Reconstructing tephra fall deposits via ensemble-based data assimilation techniques

Leonardo Mingari[1,2], Antonio Costa[3], Giovanni Macedonio[4], and Arnau Folch[2]

[1]Geosciences Barcelona (GEO3BCN-CSIC), Barcelona, Spain
[2]Barcelona Supercomputing Center, Barcelona, Spain
[3]Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna, Bologna, Italy
[4]Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Vesuviano, Naples, Italy

Abstract. In recent years, there has been a growing interest in ensemble approaches for modelling ~~volcanic plumes~~the atmospheric transport of volcanic aerosol, ash and lapilli (tephra). The development of such techniques enables the exploration of novel methods for incorporating real observations into tephra dispersal models. However, traditional data assimilation algorithms, including ensemble Kalman filter (EnKF) methods, can yield suboptimal state estimates for positive-definite variables such

5   as volcanic aerosols and tephra deposits. This study proposes two new ensemble-based data assimilation techniques for semi-positive-definite variables with highly skewed uncertainty distributions, including aerosol concentrations and tephra deposit mass loading: the GNC and GIG methods. The proposed methods are applied to reconstruct the tephra fallout deposit resulting from the 2015 Calbuco eruption using an ensemble of 256 runs performed with the FALL3D dispersal model. ~~Two~~An assessment of the methodologies is conducted considering two independent datasets of deposit thickness measurements~~are~~

10   ~~considered~~: an assimilation dataset ~~including 161 observations,~~ and a validation dataset~~for an independent assessment of the methods.~~. Different evaluation metrics (e.g. RMSE, MBE and SMAPE) are computed on the validation dataset and the results compared to two references: the ensemble prior mean and the EnKF analysis. Results show that the assimilation leads to a significant improvement over the first-guess results, obtained from the simple ensemble forecast. The evidences from this study suggest that the GNC method was the most skilful approach and represent a promising alternative for assimilation of

15   volcanic fallout data. The spatial distribution of the tephra fallout deposit thickness and ~~the ashfall~~ volume according to the ~~analyses~~GNC analysis are in good agreement with estimations based on field measurements and isopach maps reported in previous studies. ~~Both assimilation methods show a similar performance in terms of evaluation metrics and spatial distribution of the deposit. Finally, the potential application of the methodologies for the improvement of ash-cloud forecasts produced for operational models is also discussed~~On the other hand, although an interesting approach, the GIG method failed to improve

20   the EnKF analysis.

1 Introduction

Multiple hazards are associated with volcanic eruptions including lava flows, pyroclastic density currents, lahars, volcanic plumes, and tephra fallout. Specifically, the dispersal of volcanic plumes poses a serious threat to flight safety (e.g. Clarkson et al., 2016) and the subsequent fallout of tephra can cause structural damage to buildings and infrastructures due to excessive

25 loading, disrupt communication networks, airports, power plants, and water and energy distribution networks (Wilson et al., 2014). Additionally, fresh fallout deposits may be resuspended by aeolian processes affecting the air quality and prolonging the impacts of an eruption many years afterwards ~~(Folch et al., 2014; Mingari et al., 2020).~~ (Folch et al., 2014; Dominguez et al., 2020; Mingari.

The characterisation and quantification of past eruptive events is also of paramount importance for volcano hazard and risk

30 assessment studies, which infer the likelihood of future eruption scenarios based on the past volcano behaviour.

Past explosive volcanic eruptions are often characterised and classified by means of tephra deposits (Bonadonna et al., 2015) and provide critical information to infer Eruption Source Parameters (ESP) relevant to hazards, such as eruption column height, mass eruption rate, or total erupted volume (Martí et al., 2016; Constantinescu et al., 2022)~~, especially for poorly monitored volcanoes~~. Traditionally, volcanologists rely on simple field-based models to obtain certain ESP (e.g. erupted volume) assum-

35 ing an ~~exponential~~ exponential-like decay with distance for some ~~deposit parameters~~ deposit-related variables such as deposit thickness (Pyle, 1989; Bonadonna and Costa, 2013). However, it is ~~well recognised~~ well-recognised that this simplistic approach is inappropriate for tephra fall deposits with complex distribution patterns (e.g. Bonadonna et al., 1998; Martí et al., 2016). In fact, many deposits exhibit abrupt thickness variations over short distances, display well-developed secondary maxima, grain size bimodality (Durant et al., 2009), stratified deposit with alternating layer characteristics, and include other

40 complexities that make the reconstruction of tephra fallout deposits challenging (Scasso et al., 1994).

In contrast, physics-based approaches, built upon Volcanic Ash Transport and Dispersal (VATD) models, include multiple physical parameterisations and are a much more powerful tool for representing the real distribution of tephra deposits. However, the accuracy of deterministic models is highly sensitive to uncertain model input parameters (e.g. eruption column height or physical properties of particles) and the underlying meteorological fields. Alternatively, probabilistic modelling approaches

45 provide a framework to incorporate uncertainties associated with model input data. Specifically, ensemble-based modelling strategies allow one to characterise and quantify model uncertainties and have been proven to enhance VATD model skills (Bonadonna et al., 2012; Madankan et al., 2014; Stefanescu et al., 2014). For example, several VATD models have been used to conduct ensemble simulations, including ASH3D (Denlinger et al., 2012), COSMO-ART (Vogel et al., 2014), HYSPLIT (Dare et al., 2016; Zidikheri et al., 2018), NAME (Dacre and Harvey, 2018; Beckett et al., 2020), and FALL3D (Sandri et al., 2016;

50 Folch et al., 2022; Martinez et al., 2022). Furthermore, different inversion modelling techniques based on ensemble approaches have shown to produce improved volcanic ash forecasts consistent with observations by constraining ash emission estimates and model parameters (Pelley et al., 2015; Zidikheri et al., 2017; Harvey et al., 2020).

The incorporation of ensemble capabilities in VATD models set the foundations for developing and implementing ensemble-based data assimilation and inversion techniques ~~(see ?, for a recent detailed review)~~ (see Folch and Mingari, 2023, for a recent detailed rev

55 . Two main approaches have been explored in the literature to assimilate volcanic aerosol observations from satellites: ensemble Kalman filters (Fu et al., 2016, 2017; Osores et al., 2020; Pardini et al., 2020; Mingari et al., 2022) and ensemble particle filter methods (Zidikheri and Lucas, 2021a, b). Specifically, ensemble Kalman filter (EnKF) methods, used for sequential data assimilation, are based on the Kalman filter (Kalman, 1960) and approximate the probability distributions by an ensemble of system states and assume that the prior model errors and the observation noise are Gaussian. However, lower-bounded variables

such as water-vapour mixing ratio (Kliewer et al., 2016), rainfall (Husak et al., 2007) and aerosol concentrations (O'Neill et al., 2000) frequently have skewed and near-zero distributions and are not well described by Gaussian distributions. As a result, traditional EnKF methods in VATD models often yield suboptimal state estimates ~~(?)~~(Folch and Mingari, 2023).

This study explores two new ensemble-based data assimilation techniques for positive-definite variables and its implementation in VATD models, the Gaussian with Nonnegative Constraints (GNC) method and the Gamma, Inverse-Gamma, and Gaussian Ensemble Kalman Filter (GIGG-EnKF), a sequential method proposed by Bishop (2016) for highly skewed nonnegative distributions. Posselt and Bishop (2018) applied this approach for the nonlinear data assimilation of precipitation rate observations and compared the results with the analysis produced by a classical EnKF algorithm. It was concluded that the analysis ensemble of precipitation rates produced by the GIGG-EnKF bears a closer resemblance with the Bayesian posterior when the distribution is skewed.

This ~~paper~~ study aims to reconstruct the tephra fall deposit of the 2015 Calbuco eruption from a scattered set of observations. The rich existing dataset available for this eruption, consisting of deposit samples collected up to 500 km downwind from the volcano, provides an excellent test case to evaluate the proposed methodology. The Gaussian with nonnegative constraints (GNC) method and the Gamma Inverse-Gamma (GIG) method, based on the GIG equation set proposed by Bishop (2016), are used to assimilate deposit thickness data. Both methods are used here to reconstruct a complete map of the tephra-fall deposit from a dataset of uncertain observations and an ensemble of model realisations based on numerical simulations performed with the FALL3D dispersal model. In addition, a technique for emission source inversion based on the GNC method is also presented and discussed. As an initial step, this manuscript is focused on the assimilation of tephra deposits, which is crucial for long-term tephra hazard assessment, leaving to future studies the assimilation of volcanic clouds and the potential use of these two methods in operational ash forecast contexts.

The manuscript is organised as follows. The ensemble-based data assimilation methods are introduced in Sect. 2. A brief description of the 2015 Calbuco eruption is outlined in Sect. 3 where details about the observational datasets are given. Subsequently, Sect. 3 describes the numerical experiments and shows the results obtained by both methods. In Sect. 4, the GNC method is used to invert the Calbuco source term. Section 5 dwells on potential implications of the proposed methodology and possible future applications and limitations are further discussed. Finally, conclusions are drawn in Sect. 6.

2 Methodology

Data assimilation (DA) techniques have been widely used to study and forecast geophysical systems and have been applied in a variety of research and operational settings (Carrassi et al., 2018). ~~Commonly, DA~~ Data assimilation methods aim at obtaining an ~~optimal~~ estimation of the state of a dynamical system ~~by combining model forecasts with observations. Specifically, ensemble-based methods comprise two consecutive steps. In the forecast step, a prior ensemble is estimated by running an ensemble of m forward models until the observation time. In the analysis step, observational information is included to obtain a posterior estimate .~~ .(e.g. a component of the Earth system such as the atmosphere or the ocean) by exploiting information from the numerical models and the observations.

The ensemble Kalman filter (EnKF) is a remarkable example of a sequential data assimilation scheme based on the Kalman filter theory (Kalman, 1960) using a Monte Carlo approach (Evensen, 1994; Burgers et al., 1998). Given a Probability Density Function (PDF) of the model state (the so-called prior or forecast) and the observation likelihood, the goal is to estimate the updated PDF (the so-called posterior or analysis) taking into account the observation likelihood. Assume that the state of the physical system is represented by a model state vector $x \in \mathbb{R}^n$, where n is the system dimension, and that the observations are given by a vector $\tilde{y_o \in \mathbb{R}^p} y^o \in \mathbb{R}^p$, where p is the number of observations~~(see Table 1 for the list of symbols). Suppose that exists a linear observation operator $\mathbf{H} \in \mathbb{R}^{p \times n}$ which translates the model state into the observation space: .~~ EnKF uses an ensemble of model states to represent the distribution of the model state. Specifically, this ensemble-based data assimilation technique relies on a forward model which is used to generate an ensemble of trajectories of the model dynamics and the state estimate of the system is represented by an ensemble of m system state vectors $x_i \in \mathbb{R}^n$, being m the ensemble size. The average model state vector $\overline{x} \in \mathbb{R}^n$ can be approximated by the ensemble mean:

$$\tilde{= \mathbf{H}\overline{x}} \overline{x} \approx \frac{1}{m} \sum_{i=1}^{m} x_i \tag{1}$$

~~where y represent a~~ whereas the ensemble-based error covariance matrix is used to approximate the covariance $\mathbf{P}$ according to:

$$\mathbf{P} \approx \frac{1}{m-1} \mathbf{X}' \mathbf{X}'^{\mathsf{T}} \tag{2}$$

where the matrix of ensemble perturbations $\mathbf{X}' \in \mathbb{R}^{n \times m}$ is given by $[x_1 - \overline{x}, \dots, x_m - \overline{x}]$. In the EnKF framework, all PDFs are assumed to be Gaussian and, consequently, the equations (1) and (2) for mean and covariance fully define the multivariate normal distribution. The EnKF method decomposes into a forecast step and an analysis step. In the forecast step, an ensemble of model states is evolved up to the observation time using the forward model in order to estimate a prior ensemble of model states x_i^f $(i = 1 \dots m)$. No observation information is included in this forecast or prior state estimate. The analysis step is performed by updating each individual ensemble member according to

$$x_i^a = x_i^f + \mathbf{K}(y_i^o - \mathbf{H}x_i^f) \tag{3}$$

in order to generate a posterior ensemble of model states x_i^a $(i = 1 \dots m)$. The p-dimensional ~~vector. In our particular case, the model state vector x is constructed from the~~ vectors y_i^o in (3) represent an ensemble of perturbed observations with ensemble mean equal to the actual observation vector ($\overline{y_i^o} = y^o$) and error covariance $\mathbf{R}$, the observation operator is denoted by $\mathbf{H}$ and the matrix $\mathbf{K}$ is the ensemble-based Kalman gain:

$$\mathbf{K} = \mathbf{P}^f \mathbf{H}^{\mathsf{T}} (\mathbf{H} \mathbf{P}^f \mathbf{H}^{\mathsf{T}} + \mathbf{R})^{-1} \tag{4}$$

where the ensemble-based versions of the matrices $\mathbf{P}$ and measurement covariance matrix $\mathbf{R}$ are used here. Using (3), it is possible to obtain the EnKF update equations for the (ensemble-based) mean and covariance (Evensen, 1994; Burgers et al., 1998)

$$\vdots$$

$$\overline{\boldsymbol{x}}^a = \overline{\boldsymbol{x}}^f + \mathbf{K}(\boldsymbol{y}^o - \mathbf{H}\overline{\boldsymbol{x}}^f) \tag{5a}$$

$$\mathbf{P}^a = (\mathbf{1} - \mathbf{KH})\mathbf{P}^f \tag{5b}$$

125

2.1 Deposit reconstruction methods

In this work, the state of the physical system is fully determined by the two-dimensional tephra deposit load (in $\mathrm{kg\,m^{-2}}$) and the components of the model state vector $\boldsymbol{x}$ represent the mass load at the n grid points of the computational domain. ~~On the other hand, the observations vector $\boldsymbol{y}_o$~~ The FALL3D model for atmospheric passive transport and deposition of volcanic
130 tephra is used to produce the prior ensemble of volcanic deposit states $\boldsymbol{x}_i^f$ by running multiple instances of the forward model. The Eruptive Source Parameters (ESP) are perturbed around a first-guess configuration in order to define a model run for each ensemble member, as detailed in Sect. 3.3.

The observations vector $\boldsymbol{y}^o$ represents a list of scattered deposit thickness observations (in cm). ~~The~~ In consequence, the operator $\mathbf{H}$ that relates the model state to the measurements can be considered to be linear in the present case: the deposit mass
135 load (model state) and the deposit thickness at the measurement site (observation) are related by a proportionality factor, i.e. the bulk density of the tephra deposit, ~~assumed~~which, for the sake of simplicity~~for our application,~~, is assumed to be constant and equal to $800\ \mathrm{kg\,m^{-3}}$. In addition, ~~the translation of $\boldsymbol{x}$ into the observations space requires~~ horizontal bi-linear interpolations ~~for each measurement site~~are also required in order transform from the computational domain to the measurement sampling sites.

140 ~~In a~~According to the analysis scheme in the EnKF, given the prior ensemble and the observations, the deposit can be reconstructed by means of (5a) since the analysis ensemble mean can be interpreted as the best estimate (e.g. see Burgers et al., 1998) due to the underlying assumption in Kalman filters that errors are Gaussian. However, the EnKF analysis becomes suboptimal for non-Gaussian distributions. Specifically, when dealing with dispersal models of aerosols and volcanic tephra, the Gaussian assumption is critical and the EnKF analysis scheme can lead to unrealistic results in those cases, as will be shown in Sect. 3.
145 In consequence, two new methods are proposed for the reconstruction of volcanic deposits. To this purpose, let's rewrite Bayes' theorem in terms of model state mapped in the observation space. Assume that there exists a linear observation operator $\mathbf{H} \in \mathbb{R}^{p \times n}$ which translates the model state $\boldsymbol{x}$ into the observation space:

$$\boldsymbol{y} = \mathbf{H}\boldsymbol{x} \tag{6}$$

where $\boldsymbol{y}$ represent a p-dimensional vector. In a probabilistic framework, the ~~Probability Density Function (PDF)~~ PDF of the
150 state $\boldsymbol{y}$ conditioned to the observation ~~$\boldsymbol{y}_o$~~$\boldsymbol{y}^o$ is relevant to the assimilation techniques, and can be computed via the Bayes' ~~s~~ theorem (Jazwinski, 1970):

$$P(\boldsymbol{y}|\boldsymbol{y}_o^{\,o}) = \frac{P(\boldsymbol{y}^o|\boldsymbol{y})P(\boldsymbol{y})}{P(\boldsymbol{y}^o)} \tag{7}$$

where $P(\boldsymbol{y}|\boldsymbol{y}_o)$ is the a posterior PDF $P(\boldsymbol{y}|\boldsymbol{y}^o)$ is the posterior distribution, $P(\boldsymbol{y})$ is the prior PDF, $P(\boldsymbol{y}^o|\boldsymbol{y})$ is the likelihood of the data conditioned on the state $\boldsymbol{y}$ and $P(\boldsymbol{y}_o)$ are the prior PDFs, and $P(\boldsymbol{y}_o|\boldsymbol{y})$ (also known as the observation likelihood or observational PDF) is the probability of observing $\boldsymbol{y}_o$ given a true $\boldsymbol{y}$.

$P(\boldsymbol{y}^o)$ is the marginal distribution of the observation. In consequence, the determination of the posterior PDF requires the specification of both the prior and the observational PDFs. This paper proposes two ensemble-based assimilation strategies which rely on (7) and differ on the assumptions made about these PDFs. The GNC method (Sect. 2.2) uses an all-at-once assimilation approach looking for the model state that maximises the vectorial form of (7) and observations are assimilated all at once. In contrast, the GIG method (Sect. 2.3) uses a serial assimilation approach in which the univariate version of (7) is explicitly written for each single observation and the full dataset of p observations is assimilated in a sequential way.

2.2 The GNC method

The Gaussian with Nonnegative Constraints (GNC) method assumes a multi-dimensional Gaussian probability distribution for $\boldsymbol{y}$, defined in (6), given as:

$$P(\boldsymbol{y}) \propto \exp\left\{ -\frac{1}{2}(\boldsymbol{y} - \overline{\boldsymbol{y}})^{\mathsf{T}}\mathbf{P}^{-1}(\boldsymbol{y} - \overline{\boldsymbol{y}}) \right\} \tag{8}$$

where $\mathbf{P} \in \mathbb{R}^{p \times p}$ is the error covariance matrix and $\overline{\boldsymbol{y}} \in \mathbb{R}^p$ is the average model state vector in the observation space. Similarly, for the sake of simplicity, measurements are assumed to be normally distributed with observation error covariance matrix $\mathbf{R} \in \mathbb{R}^{p \times p}$, i.e.

$$P(\boldsymbol{y}_o{}^o|\boldsymbol{y}) \propto \exp\left\{ -\frac{1}{2}(\boldsymbol{y}_o{}^o - \boldsymbol{y})^{\mathsf{T}}\mathbf{R}^{-1}(\boldsymbol{y}_o{}^o - \boldsymbol{y}) \right\} \tag{9}$$

The most likely state is the one that maximises the posterior PDF, Eq. (7), or equivalently the one that minimises the GNC cost function J:

$$J(\boldsymbol{y}) \propto (\boldsymbol{y} - \overline{\boldsymbol{y}})^{\mathsf{T}}\mathbf{P}^{-1}(\boldsymbol{y} - \overline{\boldsymbol{y}}) + (\boldsymbol{y}_o{}^o - \boldsymbol{y})^{\mathsf{T}}\mathbf{R}^{-1}(\boldsymbol{y}_o{}^o - \boldsymbol{y}) \tag{10}$$

Note that the expression above is actually very similar to the cost function used in classical variational methods (e.g. 3DVAR, Carrassi et al., 2018) with the difference that $\overline{\boldsymbol{y}}$ plays the role of the model background state in the VAR methods, and the first term in (10) is computed in the observations space rather than in the model space as usual. This is justified because expressing the functional J in the observations space is advantageous in those cases where observations are localised and/or nearly zero, i.e. circumscribed to portion of the computational domain (this is what typically occurs for volcanic clouds and fallout deposits). Moreover, the functional in (10) yields to a much reduced system when compared to classical VAR methods because the observation space normally has a much lower dimension ($p << n$).

In the context of an ensemble modelling formulation, the forecast step provides an ensemble Given a prior ensemble of m independent prior states vectors $\boldsymbol{x}_i$ representing m model realisations at the analysis time. The , the GNC method looks for the best linear estimate of the system state in the subspace spanned by the ensemble of vectors $\boldsymbol{x}_i$:

$$\boldsymbol{x} = w_1\boldsymbol{x}_1 + \cdots + w_m\boldsymbol{x}_m \tag{11}$$

Table 1. List of symbols using the following convention: matrices in upper case bold, vectors in lower case bold, scalars in italics.

Symbol	Description
General definitions	
m	ensemble size
n	dimension of model state vector
p	number of observations
$\boldsymbol{x} \in \mathbb{R}^n$	Model state vector
$\boldsymbol{y}^o \in \mathbb{R}^p$	Observations vector
$\boldsymbol{y} \in \mathbb{R}^p$	Model state vector in the observation space
$\mathbf{H} \in \mathbb{R}^{p \times n}$	Observation operator
GNC method	
$\mathbf{P} \in \mathbb{R}^{p \times p}$	Model covariance matrix
$\mathbf{R} \in \mathbb{R}^{p \times p}$	Observation error covariance matrix
$\boldsymbol{w} \in \mathbb{R}^m$	Vector of weight factors
$\overline{\boldsymbol{y}} \in \mathbb{R}^p$	Average model state vector (obs. space)
$\mathbf{Y} \in \mathbb{R}^{p \times m}$	Ensemble model state matrix (obs. space)
$\mathbf{Y}' \in \mathbb{R}^{p \times m}$	Ensemble perturbations matrix (obs. space)
GIG method	
y_j	j-th component of $\boldsymbol{y}$
y_j^o	j-th component of $\boldsymbol{y}^o$
$\overline{y}_j^f$	Mean of prior distribution of y_j
$\overline{y}_j^a$	Mean of analysis distribution of y_j (19a)
P_j^r	Type 1 relative error variance of prior $P_j^r := \mathrm{var}(y_j^f)/(\overline{y}_j^f)^2$
Π_j^r	Type 1 relative error variance of analysis (19b) $\Pi_j^r := \mathrm{var}(y_j^a)/(\overline{y}_j^a)^2$
R_j^r	Type 1 relative error variance of observation[†] $R_j^r := \mathrm{var}(y_j^o)/(y_j^t)^2$
$\tilde{P}_j^r$	Type 2 relative error variance of prior $(\tilde{P}_j^r)^{-1} = (P_j^r)^{-1} + 1$
$\tilde{R}_j^r$	Type 2 relative error variance of observation $(\tilde{R}_j^r)^{-1} = (R_j^r)^{-1} + 1$

[†] where y_j^t is the true value of the j-th observation.

where $w_i \geq 0 \, (i = 1 \ldots m)$ is a set of nonnegative weight factors for each ensemble member. The important point here is that the nonnegative constraints on w_i relax the Gaussian hypotheses and avoid the occurrence of non-physical solutions. The linearity of the observation operator $\mathbf{H}$ allows the analysis to be expressed in the observation space as:

$$y = w_1 y_1 + \cdots + w_m y_m = \mathbf{Y} w \tag{12}$$

where $y_i = \mathbf{H} x_i$ and the matrix $\mathbf{Y} \in \mathbb{R}^{p \times m}$ is defined by $[y_1, \ldots, y_m]$. The ensemble mean is used to approximate the average model state vector $\overline{y} \in \mathbb{R}^p$, i.e.:

$$\overline{y} \approx \frac{1}{m} \sum_{i=1}^{m} y_i \tag{13}$$

whereas the ensemble-based error covariance matrix is used to approximate $\mathbf{P}$ according to:

$$\mathbf{P} \approx \frac{1}{m-1} \mathbf{Y}' \mathbf{Y}'^\mathsf{T} \tag{14}$$

where the matrix of ensemble perturbations $\mathbf{Y}' \in \mathbb{R}^{p \times m}$ is given by $[y_1 - \overline{y}, \ldots, y_m - \overline{y}]$. Replacing (12), (13) and (14) in (10), the GNC cost function J can be expressed as the equivalent quadratic form:

$$J(w) = \frac{1}{2} w^\mathsf{T} \mathbf{Q} w + b^\mathsf{T} w + \ldots \tag{15}$$

with

$$\mathbf{Q} = \mathbf{Y}^\mathsf{T} (\mathbf{P}^{-1} + \mathbf{R}^{-1}) \mathbf{Y} \tag{16a}$$

$$b = -\mathbf{Y}^\mathsf{T} (\mathbf{P}^{-1} \overline{y} + \mathbf{R}^{-1} y_o{}^o) \tag{16b}$$

In order to find the optimal vector of weight factors w, the optimisation problem $\min_{w \geq 0} J(w)$ must be solved. Then, the analysis vector state x^a can be computed by replacing the optimal w in (11). The minimisation of the quadratic form (15) subject to the constraints $w_i \geq 0 \, \forall \, i$ is a nonnegative quadratic programming problem and there is no analytical solution for the global minimum due to the nonnegativity constraint. However, it can be solved using the iterative approach proposed by Sha et al. (2007):

$$w_i \leftarrow w_i \left[\frac{-b_i + \sqrt{b_i^2 + a_i c_i}}{a_i} \right] \tag{17}$$

as long as the matrix $\mathbf{Q}$ is symmetric and semipositive definite, as can be easily verified from (16a). Under the multiplicative updates (17), the cost function decreases monotonically to the value of its global minimum as shown by Sha et al. (2007). The vectors $a = \mathbf{A}_+ w$ and $c = \mathbf{A}_- w$ must be updated in each iterative step, where $\mathbf{A}_+ = |\mathbf{Q}| + \mathbf{Q}$ and $\mathbf{A}_- = |\mathbf{Q}| - \mathbf{Q}$. For illustrative purposes, Fig. 1 shows the decrease of the normalised cost function, defined as $\sqrt{J/p}$, under the multiplicative updates in (17) for the case study presented later in Sect. 3. More than 10^4 iterative steps were required to get low enough residuals to satisfy the convergence criteria.
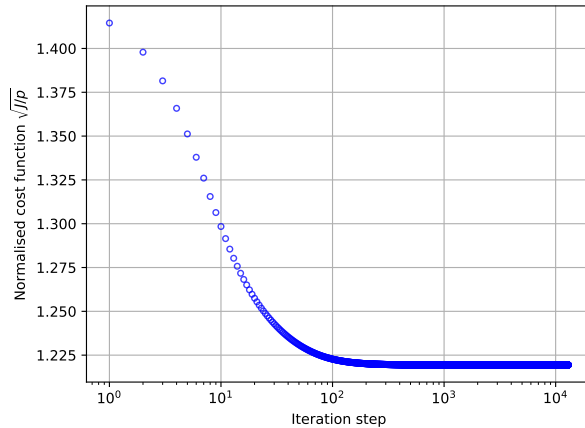
Figure 1. Iterative approach to minimise the GNC cost function J subject to the non-negativity constraints. Under the multiplicative updates in (17) the cost function decreases monotonically. In this particular example, the convergence required more than 10^4 iterative steps.

2.3 The GIG method

Bishop (2016) introduced a variation of the Ensemble Kalman Filter (EnKF) that solves the univariate Bayes' theorem for non-negative variables with skewed (asymmetrical) probability distributions. The so-called GIGG-EnKF (with GIGG standing for Gamma, Inverse-Gamma and Gaussian) allows non-negative variables typically involving near-zero values (i.e. with right-skewed probability distributions), such as aerosol, water vapour, cloud, and precipitation concentrations, to be directly assimilated, thus avoiding the use of Gaussian anamorphosis nonlinear transformations (e.g. Amezcua and Leeuwen, 2014). The GIGG-EnKF algorithm is based on the generalised two-stage multivariate ensemble filter described by Anderson (2003). The first stage involves the univariate GIGG-EnKF in which an ensemble-based estimate of the posterior distribution of the observed variable is generated from a single observation and a prior ensemble of state estimates. In the second stage, the univariate method is extended by propagating this information to the complete model state vector using a linear regression approximation.

According to the strategy proposed by Bishop (2016), in the first step, the Bayes' theorem is solved in the univariate form of (7) assuming a distribution pair for the prior probability and the likelihood PDF of the error-prone observations given truth, respectively. A single observation is assimilated using an appropriate equation set depending on three different cases: GIG, IGG and G. ~~This paper will only consider the~~ The GIG equation set ~~,~~ is aimed at situations in which the prior can be described by a Gamma distribution and the observation likelihood can be represented by an Inverse-Gamma distribution. In addition, Bishop (2016) introduced the IGG equation set (Inverse-Gamma prior and Gamma observation likelihood) and the G equation set (Gaussian distributions).

Volcanic aerosols have been found to be well approximated by gamma distributions (Mingari et al., 2022). Similarly, it is shown in this paper that the prior distributions of deposit mass loading can be well represented to some extent by gamma

distributions (see Sect. 3.4). Consequently, this paper will focus exclusively on the GIG case. In the GIG case, the posterior probability is given by a gamma PDF:

$$P(y_j|y_j^o) \propto y_j^{(\Pi_j^r)^{-1}-1} \exp\left\{-\frac{y_j}{\Pi_j^r\,\overline{y}_j^a}\right\} \tag{18}$$

where y_j and y_j^o are the j-th components of the vectors $\boldsymbol{y}$ and ~~$\boldsymbol{y}_o$~~ $\boldsymbol{y}^o$ respectively. The posterior univariate gamma PDF is characterised by two parameters, namely, the analysis mean $\overline{y}_j^a$ and the type 1 relative error variance of the analysis $\Pi_j^r :=$ $\mathrm{var}(y_j^a)/(\overline{y}_j^a)^2$, that in the GIG method are given by:

$$\frac{1}{\overline{y}_j^a} = \frac{1}{\overline{y}_j^f} + \frac{\tilde{P}_j^r}{\tilde{R}_j^r+\tilde{P}_j^r}\left\{\frac{1}{y_j^o} - (\tilde{R}_j^r+1)\frac{1}{\overline{y}_j^f}\right\} \tag{19a}$$

$$(\Pi_j^r)^{-1} = (\tilde{R}_j^r)^{-1} + (\tilde{P}_j^r)^{-1} \tag{19b}$$

where $\tilde{P}_j^r$ and $\tilde{R}_j^r$ are the type 2 relative error variance of the prior and observations respectively (see Table 1 for details). ~~Bishop (2016) provides a stochastic equation to generate individual members of the analysis ensemble (y_{ji}^a)~~

In order to generate an analysis ensemble y_{ji}^a with the low-order moments of the posterior distribution being consistent with (19)~~. See Appendix ?? and~~, Bishop (2016) provides the following stochastic equation for the case of an univariate gamma prior and an inverse-gamma observation-likelihood PDFs:

$$\frac{y_{ji}^a-\overline{y}_j^a}{\overline{y}_j^a} = \frac{y_{ji}^f-\overline{y}_j^f}{\sqrt{\left(\overline{y}_j^f\right)^2 + \mathrm{var}(y_{ji}^f)}} + \tilde{P}_j^r\left(\tilde{P}_j^r+\tilde{R}_j^r\right)^{-1} \times$$
$$\times \left\{\frac{z_{ji}-\overline{z}_j}{\sqrt{\left(\overline{z}_j\right)^2 - 2\mathrm{var}(z_{ji})}} - \frac{y_{ji}^f-\overline{y}_j^f}{\sqrt{\left(\overline{y}_j^f\right)^2 + \mathrm{var}(y_{ji}^f)}}\right\} \tag{20}$$

where $\overline{y}_j^a$ can be computed using Eq. (~~20) for further details.~~19a). Equation (20) ensures that the analysis ensemble y_{ji}^a is consistent with the type 1 relative error variance of the posterior given by Eq. (19b) provided z_{ji} is randomly sampled from a gamma PDF with type 1 relative error variance R_j^z and mean $\overline{z}_j$ given by:

$$\left(R_j^z\right)^{-1} = \left(\tilde{R}_j^r\right)^{-1} + 2 \tag{21a}$$

$$\overline{z}_j = \frac{\left(\tilde{R}_j^r\right)^{-1}+2}{\left(\tilde{R}_j^r\right)^{-1}}y_j^o \tag{21b}$$

In addition, the ensemble generated in this way is ensured to converge to the true posterior PDF for large ensembles.

~~In the second step, the~~ The univariate case is extended according to the second-stage linear regression step proposed by Anderson (2003) in order to find the analysis ensemble for the complete model state vector. The update of the k-th model state

vector variable of the i-th ensemble member due to the j-th observation is computed according to:

$$x_{ki}^a = x_{ki}^f + \delta_{kj} \, \frac{\text{cov}(x_k^f, y_j^f)}{\text{var}(y_j^f)} (y_{ji}^a - y_{ji}^f) \tag{22}$$

255 ~~where δ_{kj} is a localisation weight satisfying $0 \leq \delta_{kj} \leq 1$. In the original formulation, Bishop (2016) assumed $\delta_{kj} = 1$. However, we found that (22) can introduce artificial results when a poor correlation exist between the x_k and y_j variables, including non-physical values (negative mass load). By introducing the localisation weight δ_{kj} in this work, the effect of poorly correlated observations is intended to be limited. Instead of the traditional distance-dependent localisation, we use here the correlation-cutoff method to localise the analysis (Yoshida and Kalnay, 2018). This method localises the observation impacts based on their~~

260 ~~background error correlations and only cross covariances between variables that have strong background error correlation are considered in (22). In this work, a quadratic function has been chosen for defining the cutoff function (see Chang and Kalnay, 2022, for a sim . Specifically, the localisation weight assigned for the j-th observation at the k-th analysis grid cell was defined as:~~

$$\delta_{kj} = \begin{cases} 1 & \text{if } |\rho_{kj}| \geq \rho_o \\ (\rho_{kj}/\rho_o)^2 & \text{if } |\rho_{kj}| < \rho_o \end{cases}$$

~~where $\rho_{kj} = \text{corr}(x_k^f, y_j^f)$ is the prior ensemble correlation and ρ_o is a tunable parameter controlling the intensity of the~~

265 ~~localisation. Localisation is enabled when $0 < \rho_o < 1$ and $\rho_o = 0$ implies no localisation at all, as in the original formulation.~~

The inverse-gamma PDFs assign non-zero probability densities only for positive observations and, as a result, zero observations cannot be properly assimilated using the GIG equation set (e.g. see 19a). This problem is addressed here by redefining zero observation data according to:

$$y_j^o \leftarrow r * \epsilon_{min} \tag{23}$$

270 where $r \in (0, 1]$ is a random number and ϵ_{min} is ~~the minimum~~ a typical error expected for ~~the deposit thickness observations~~ zero-valued observations of deposit thickness (assumed to be $\epsilon_{min} = 1$ mm in this work ~~where we deal only with~~ since only visible tephra deposits are considered here).

The GIG method is a sequential procedure: a single observation is assimilated in order to update the prior ensemble forecast using the GIG equation set; subsequently, this procedure is repeated until all observations have been sequentially assimilated. In

275 contrast, the GNC method described in Sec. 2.2 represents an all-at-once assimilation technique. Another important difference is that the GIG method is stochastic, i.e.different applications of the method will lead to different realisations of the analysis. To summarise, a pseudocode of the sequential procedure used to implement the GIG method is detailed in the Algorithm 1.

3 Reconstruction of the 2015 Calbuco deposit

In this section, the procedures described in Sect. 2 are applied to the 2015 Calbuco eruption in order to obtain the analysed de-

280 posit thickness. With this in mind, ~~the field measurements reported by Van Eaton et al. (2016)~~ a total of 204 field measurements

Algorithm 1 Pseudocode of the GIG method based on the Bishop (2016) algorithm for the case in which the prior is a gamma distribution and the observation likelihood is an inverse-gamma distribution.

Require: List of observations $\{y_j^o\}$ with their relative errors
Ensure: Analysis ensemble x_{ki}^a

0: **get** x_{ki}^f {Generate the prior ensemble}
0: **procedure** GIG METHOD
0: randomly shuffle observation list
0: **for all** y_j^o **do** {Iterate over observations}
0:    **if** $y_j^o = 0$ **then**
0:      $r \leftarrow$ random number $\in (0, 1]$
0:      $y_j^o \leftarrow r * \epsilon_{min}$ {Redefine zero observations}
0:    **end if**
0:    $\overline{y}_j^a \leftarrow$ Eq. (19a)
0:    $y_{ji}^a \leftarrow$ Eq. (20)
0:    $x_{ki}^a \leftarrow$ Eq. (22) {Generate the analysis ensemble}
0:    $x_{ki}^f \leftarrow x_{ki}^a$ {Update the prior ensemble}
0: **end for**
0: **end procedure**=0

---

of deposit thickness will be considered for assimilation ~~purposes.~~ and validation purposes. This dataset is composed by 160 measurements reported by Van Eaton et al. (2016) and an independent dataset of 44 measurements provided by F. Reckziegel (pers. comm., September, 2020). Figure 2 shows the location of the sampling sites for both datasets.

3.1 Fallout deposit and datasets

285 The 2015 eruption of the Calbuco stratovolcano (41.33°S, 72.61°W) in Southern Chile involved two major eruptive pulses on 22–23 April along with a third minor pulse on 30 April (Romero et al., 2016). During the most energetic phase on 23 April, stratospheric eruption columns higher than 15 km above the vent level ($\sim$17 km above sea level) were developed. Regions over Southern Chile and the Argentinian Patagonia were severely affected by tephra fall. According to different estimations based on field studies, deposit volume ranges between 0.27 and 0.58 km^3 (Romero et al., 2016; Van Eaton et al., 2016).

290 The availability of independent and comprehensive datasets of field observations makes the Calbuco tephra deposit an excellent case study. Van Eaton et al. (2016) reported the thickness and stratigraphy of the fall deposits at 163 sampling sites within a 500-km radius from the volcano summit. ~~This dataset is considered in this work for assimilation purposes.Figure 2shows the location of sampling sites and the isopachs of fall deposit thickness (in centimetres) reported by Van Eaton et al. (2016)~~ In addition, a complementary dataset composed of 45 independent measurements of deposit thickness (F. Reckziegel, pers.

295 comm.) distributed over a larger region is also available (Fig. 2). Finally, a hand-drawn isopachs map built from a third
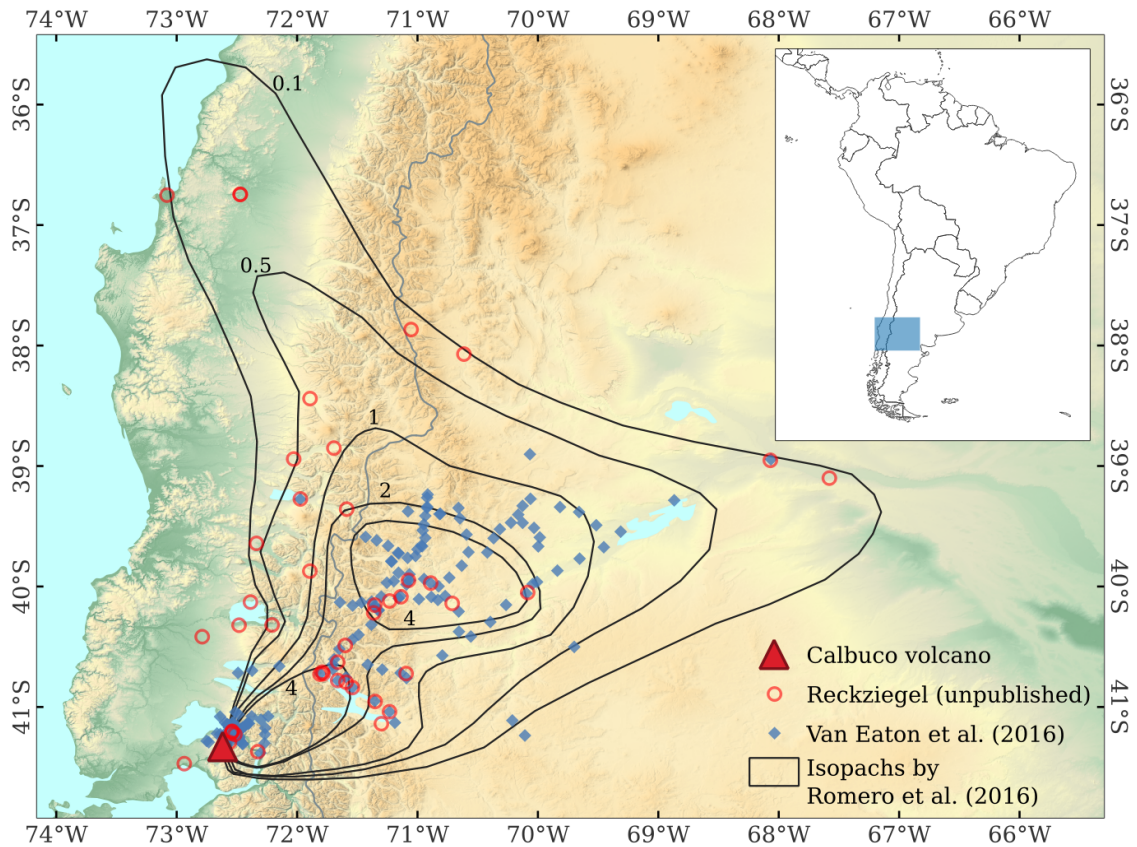
Figure 2. Location of the sampling sites corresponding to the dataset reported by Van Eaton et al. (2016) (blue diamond) and an independent dataset composed by 45 measurements (red circle) provided by F. Reckziegel (pers. comm., September, 2020). The map also shows the isopachs of fall deposit thickness in millimetres reported by Romero et al. (2016) used for validation purposes in this work.

independent dataset (Romero et al., 2016) is also used to evaluate the tephra deposit distribution. The corresponding isopachs for 0.1, 0.5, 1.0, 2.0 and 4 mm are represented in Fig. 2. The three datasets are summarised in Table 2.

It is interesting to note the presence of a secondary thickness maximum ∼200 km downwind from the vent, located around two major cities of the Argentinian Patagonia: Junín de los Andes (39.95°S, 71.07°W) and San Martín de los Andes (40.16°S, 71.35°W), likely due to ash aggregation processes (e.g. Costa et al., 2010). The emergence of this distal maximum indicates that a complex plume dynamics was involved in the volcanic eruption.

Isopachs of fall deposit thickness in centimetres and location of the 161 sampling sites (two ambiguous measurements were discarded) of the assimilation dataset reported by Van Eaton et al. (2016) (red stars). The map also shows the locations of an independent dataset composed by 45 measurements (blue squares) used for validation purposes.

A complementary dataset composed of 45 independent measurements of deposit thickness (F. Reckziegel, pers. comm.) is also considered for validation purposes using the evaluation metrics described in the following section. Figure 2 shows

Table 2. Calbuco deposit datasets considered in this study.

Reference	Data type	Purpose
Van Eaton et al. (2016)	Thickness at 163 locations	Data assimilation
Reckziegel (unpublished)	Thickness at 45 locations	Data validation
Romero et al. (2016)	Isopachs map	Data validation

~~the location of the corresponding sampling sites (square symbols). Finally, a hand-drawn isopachs map built from a third independent dataset (Romero et al., 2016) is also used to evaluate the tephra deposit distribution considering the isopachs for 0.1, 0.5, 1.0 and 2.0 mm. The three datasets are summarised in Table 2.~~ The assimilation methods require a dataset of measurements along with the corresponding absolute or relative errors. Specifically, the GNC method requires the absolute error ϵ_j associated with the j-th measurement y_j^o (the observation error covariance matrix $\mathbf{R}$ is assumed diagonal with elements ϵ_j^2). On the other hand, the GIG method requires the relative error $\epsilon_j^r = \epsilon_j / y_j^t$, where y_j^t is the true value of the j-th observation.

~~Observation error~~ The strategy adopted in this work to provide reasonable error estimates is based on a clustering algorithm and observation error standard deviations are assumed to be dependent on the measured value. ~~The procedure for determining observation errors is as follows: the full dataset of observations is classified in 7 groups using a spectral clustering algorithm (Pedregosa et al., 2011), as shown in Fig. ??. Three feature were used for clustering: latitude, longitude and deposit thickness of every measurement~~ In summary, observational data are organized into groups with similar characteristics and an absolute and relative error is assigned to each group or cluster. The error for the j-th measurement y_j^o is approximated by standard deviation associated with the corresponding cluster; the true value y_j^t, required to estimate relative errors, is approximated by the cluster mean value. ~~Observational data are classified in 7 groups using a spectral clustering algorithm taking into account three features: latitude, longitude, and deposit thickness of each measurement. The observation absolute error is assumed to be the standard deviation for the corresponding cluster and the relative error is approximated by the ratio of standard deviation to the cluster mean.~~ A more detailed explanation of the strategy used to estimate errors can be found in the supplementary material.

3.2 Validation metrics

As validation metrics, we consider the weighted versions of the Mean Bias Error (MBE) and the Root-Mean-Square Error (RMSE) to measure the differences between observations and analyses, defined as~~usual~~:

$$\text{MBE} = \frac{1}{p} \sum_{j=1}^{p} \omega_i (y_j^o - y_j) \tag{24a}$$

$$\text{RMSE} = \sqrt{\frac{1}{p}\sum_{j=1}^{p}\omega_i^2(y_j^o - y_j)^2} \tag{24b}$$

330 ~~The~~ Notice that if the non-weighted ($\omega_j = 1$) versions of (24) are used, the MBE (in cm) quantifies the tendency to overestimate or underestimate observations for the overall dataset whereas the RMSE (in cm) measures the average magnitude of the errors. These two metrics are suitable when a uniform distribution of the errors is expected. However, the datasets in this work contain measurements spanning four orders of magnitude and the assumption of a constant absolute error seems to be inappropriate in this case because only proximal data (i.e. the largest measurements of deposit thickness) contributes significantly to the

335 non-weighted versions of MBE and RMSE. ~~The weighted MBE and RMSE represent alternative evaluation metrics that are more meaningful in our case. These dimensionless metrics are defined according to :~~

$$\text{wMBE} = \frac{1}{p}\sum_{j=1}^{p}\frac{y_j^o - y_j}{\epsilon_j}$$

$$\text{wRMSE} = \sqrt{\frac{1}{p}\sum_{j=1}^{p}\frac{(y_j^o - y_j)^2}{\epsilon_j^2}}$$

~~where the deviations from observation ($y_j^o - y_j$) are weighted by~~ Instead, in order to treat the deviations more evenly, the

340 observation uncertainties ~~(ϵ_j). The~~ are used to define the weights according to $\omega_j = 1/\epsilon_j$. This weighted versions of MBE and RMSE represent dimensionless evaluation metrics and the impact of the assimilation can be better characterised by means of these metrics. Ideally, the ~~wMBE~~ MBE should be close to zero and ~~wRMSE should approach one.~~ RMSE should be close to one.

Another relative metric used to evaluate the deposit reconstruction is the Symmetric Mean Absolute Percentage Error

345 (SMAPE), defined as:

$$\text{SMAPE} = 100 \times \frac{2}{p}\sum_{j=1}^{p}\frac{|y_j^o - y_j|}{|y_j^o| + |y_j|} \tag{25}$$

and is expressed in percentages. Notice that this metric does not depend on the observation errors.

3.3 Ensemble modelling

Numerical simulations were carried out using the latest version release of FALL3D (v8.2), an open-source offline Eulerian

350 model for atmospheric ~~passive~~ transport and deposition of aerosols and particles, including tephra species. FALL3D solves the so-called Advection-Diffusion-Sedimentation (ADS) equation (Folch et al., 2020; Prata et al., 2021). The new FALL3D version has been designed to support increasingly larger scientific workloads and prepare the code for the transition to extreme-scale computing systems (Folch et al., 2023). Specifically, the code version v8.x has been released with several improvements over previous versions, including improvements in the model physics, numerical algorithmic methods, and computational efficiency.

355 In addition, from version v8.1 onwards, the FALL3D model enables ensemble simulations to be performed very efficiently

Table 3. FALL3D model configuration parameters for the 2015 Calbuco runs.

Parameter	Value
Ensemble size	256
Resolution	$0.05° \times 0.05°$
Number of grid points	$180 \times 160 \times 45$
Species	40 tephra bins
Grain Size Distribution	bi-Gaussian
Run time	100 h
Emission source	Suzuki source (Pfeiffer et al., 2005)
Mass emission rate	Estimated from column height (Degruyter and Bonadonna, 2012)

by means of a single parallel task (Folch et al., 2022). Ensemble modelling allows one to characterise and quantify model uncertainties due to poorly constrained input parameters and errors in the model physics parameterisations or the underlying model-driving meteorological data. In addition, the ability to generate ensemble runs makes it possible to improve forecasts by incorporating observations using different ensemble-based data assimilation techniques.

360   The configuration of the FALL3D model used in this work is summarised in Table 3. A three-dimensional computational domain with a horizontal resolution of $\sim$4 km ($0.05°$) and $180 \times 160 \times 45$ grid points was defined. The Total Grain-Size Distribution (TGSD) of tephra injected into the atmosphere consists of the sum of two log-normal distributions (i.e. bi-Gaussian in Φ-units) including 40 tephra bins. The modes and standard deviations of the bimodal distribution were computed using the parameterisation proposed by Costa et al. (2016), which estimates them from the eruption intensity and magma viscosity. The

365   mode of the coarser population was located at -1.2Φ with a standard deviation of 1.71Φ, while the mode of the finer population was 3.49Φ with a standard deviation of 1.46Φ. The weight of each subpopulation was set to $p_c = 0.15$ and $p_f = 0.85$ for the coarse and fine population, respectively. The vertical mass distribution of the source term depends on the eruptive column top height (H) according to the following parameterisation (Pfeiffer et al., 2005):

$$\frac{dm}{dz} \propto \left\{ \left(1 - \frac{z}{H}\right) \exp\left[A_s\left(\frac{z}{H} - 1\right)\right] \right\}^{\lambda_s} \tag{26}$$

370   where A_s and λ_s are the so-called Suzuki parameters (Suzuki, 1983; Pfeiffer et al., 2005). Finally, the mass emission rate was computed from the eruptive column top height using the ~~expression~~ relationship proposed by Degruyter and Bonadonna (2012).

~~An ensemble size of 256 members is considered in this paper. The~~ A 256-member prior ensemble was generated by perturbing the Eruption Source Parameters (ESP) and the horizontal wind components around a reference value using ei-

375   ther uniform or truncated normal distributions ~~with the range of parameter uncertainties. The Latin Hypercube Sampling (LHS, McKay et al., 1979) was used to efficiently sample the parameter space.~~. Table 4 lists the perturbed model parameters along with the corresponding reference values and sampling uncertainty ranges. The ~~central member of the ensemble was~~

Table 4. Ensemble configuration. The perturbed model parameters are: eruption column height (H), eruption phase start time (T_i), phase duration (ΔT), parameters A_s and λ_s of the Suzuki vertical mass distribution, the fine mode of the bi-Gaussian TGSD, the density of aggregates, and the wind components.

Parameter	Reference value	Distribution	Sampling range
H	15 & 15 km avl[†]	Uniform	$\pm 25\%$
T_i	21 & 28 h[‡]	Uniform	15 min
ΔT	1.5 & 6 h	Uniform	15 min
A_s	6	Uniform	3
λ_s	3	Uniform	2
Fine mode	3.49 Φ	Uniform	2 Φ
Density agg.	450 kg m^{-3}	Uniform	100 kg m^{-3}
U wind	ERA5[§]	Gaussian	$\pm 20\%$
V wind	ERA5[§]	Gaussian	$\pm 20\%$

[†] Two eruptive phases are considered. Heights are given in km above the vent level;

[‡] In hours since 22 April 2015 at 00:00 UTC;

[§] ECMWF atmospheric reanalysis (137 model levels).

~~defined considering two eruptive phases with~~ Latin Hypercube Sampling (LHS, McKay et al., 1979) was used to efficiently sample the parameter space. Two eruptive phases were considered with reference values of column top heights at $H =$
380 15 km avl (above vent level) for each phase. The ~~ensemble was generated by perturbing~~ column top height H was independently perturbed for each phase ~~independently and~~ assuming a sampling range of $\pm 25\%$. The source start time for the central member was defined at 21:00 UTC on 22 April 2015 (1st phase) and at 04:00 UTC on 23 April 2015 (2nd phase) assuming a duration of 1.6 h and 6 h for each phase. Source start times and phase duration were also perturbed.

3.4 Prior ensemble distribution

385 Before showing how the assimilation methods perform, it is worth to ~~consider the prior (forecast)~~ characterise the prior ensemble distribution and to check weather the ~~PDF $P(\boldsymbol{y})$ fulfils the assumption~~ assumptions of the GIG method are fulfilled (i.e. ~~a gamma prior distribution~~ the prior distribution can be approximated by a gamma PDF). To this purpose, the skewness $\tilde{\mu}_3$ (i.e. the third standardised moment, μ_3/σ^3) of the prior distribution was computed from the random samples y^f_{ji}, i.e. the forecasted deposit thickness according to the i-th ensemble member ($i = 1\dots m = 256$ $i = 1\dots m$) at the sampling site of the j-th obser-
390 vation ($j = 1\dots p = 161$ $j = 1\dots p$). Figure 3 shows the results for each observation point by plotting $\tilde{\mu}_3$ as a function of the standard deviation-to-mean ratio, i.e. $\sqrt{P^r_j}$ ~~.~~ (using the notation introduced in Sect. 2.3). For illustrative purposes, the skewness of three different theoretical distributions (Gaussian, log-normal and gamma) is shown. As expected, the symmetric Gaussian distribution, characterised by $\tilde{\mu}_3 = 0$, does not reproduce the positively skewed prior distribution. The log-normal family of probability distributions represent an example of distributions for positive-definite variables with a lower bound. However, as
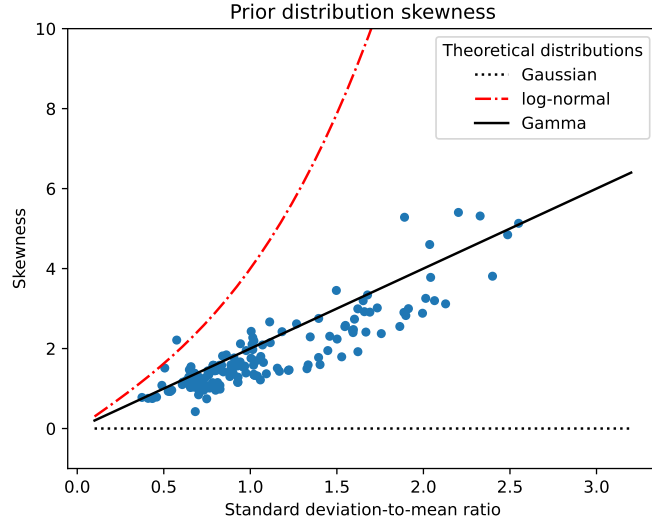
Figure 3. Skewness as a function of the standard deviation-to-mean ratio for the prior distribution at the sampling locations (blue circles). Results for some theoretical distributions (Gaussian, log-normal and gamma) are also shown for comparison.

395 shown in Fig. 3, the log-normal distribution cannot properly represent the prior distribution because the theoretical skewness is extremely large in this case. In contrast, the skewness computed from the prior distributions (blue dots) is well approximated by the relationship $\tilde{\mu}_3 = 2\sqrt{P_j^r}$, which is the theoretical expression corresponding to the gamma distribution (solid black line).

In order to ~~dig further into~~ further understand the similarities between the gamma and prior distributions, Fig. 4 explicitly shows histograms of sampled prior distributions along with the corresponding theoretical gamma distribution for some obser-

400 vation sites. The theoretical gamma distributions were constructed using the sampled first and second moments. As observed, a good agreement is found between both distributions in almost all cases. Note that when the type 1 relative error variance is greater than 1 ($P_j^r > 1$) the gamma probability density decreases monotonically (Fig. 4~~a–b~~a–c) and the mode becomes zero. In contrast, when $P_j^r < 1$ or, equivalently, when $\overline{y}_j^2 > \mathrm{var}(y_j)$, the mode becomes positive (Fig. 4~~c–p~~d–p). The results obtained is this section justify the suitability of the GIG method to deal with the assimilation volcanic deposit data.

405 ## 3.5 Analyses

~~The GNC method~~ In this section, we compare the tephra deposit field reconstructed according to 4 strategies: (i) Forecast, i.e.the prior ensemble mean, (ii) EnKF method, i.e.the analysis ensemble mean via (5a), (iii) GNC method and (iv) GIG method. Notice that the GNC method gives a set of weight factors for each ensemble member ~~(w_i ,~~($i = 1, \ldots, m$) and the best estimate of the system state is obtained by replacing the optimal weight factors in (11). On the other hand, the GIG method produces

410 an ensemble of analysis states ~~. In this section, and~~ the analysis ensemble mean ~~corresponding to the GIG method~~ is used for comparison purposes in this section. The full dataset of 204 measurements were assimilated to compute the analysis according to the strategies (ii), (iii) and (iv).
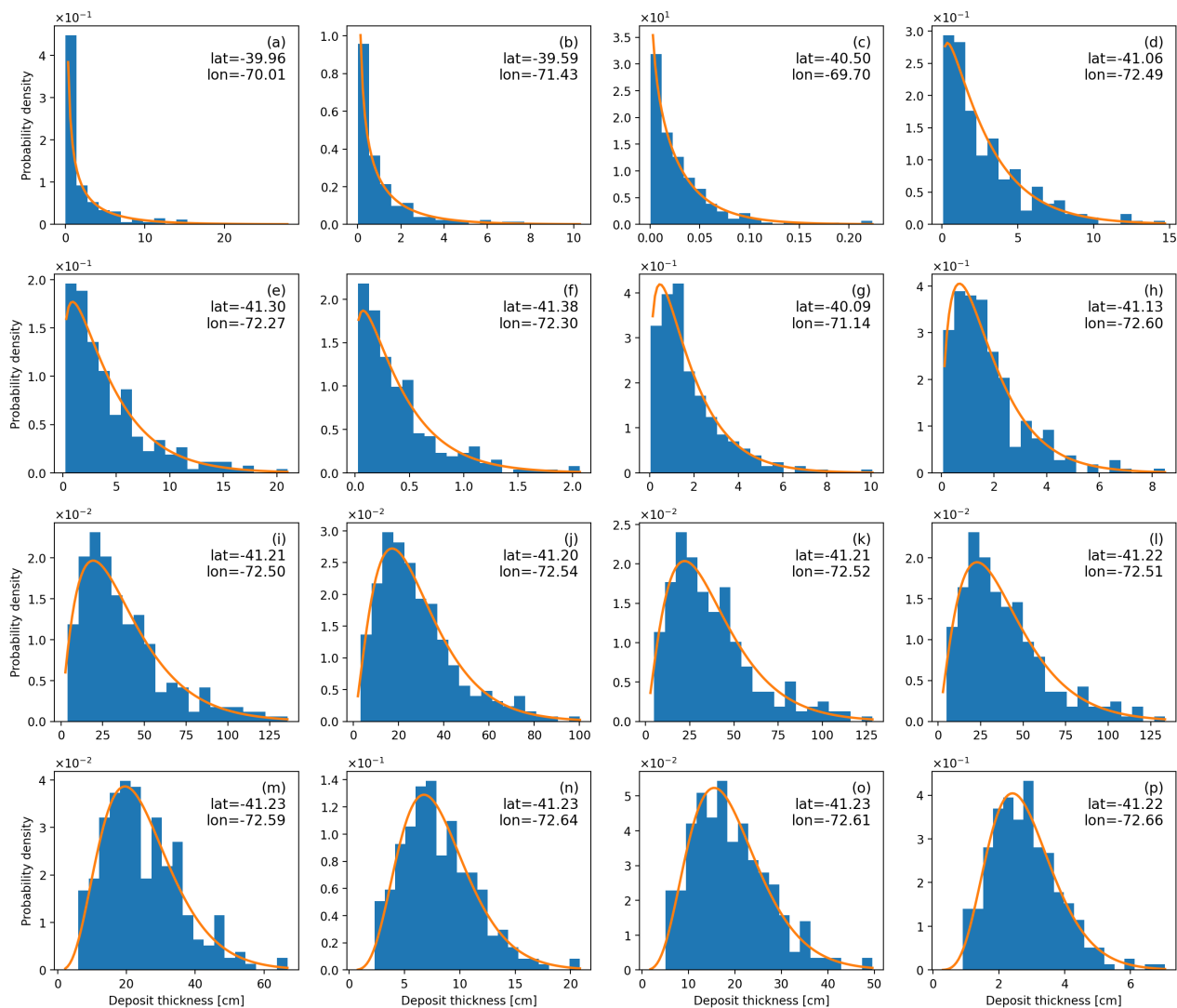
Figure 4. Histograms of sampled prior distributions along with the corresponding theoretical gamma distribution (solid line) at some selected observation sites.

The results of the tephra fall deposit reconstruction are shown in Fig. 5. For comparative purposes, the Romero et al. (2016) deposit contours (hand-drawn isopachs for 0.1, 0.5, 1, 2 and 4 mm) are over imposed in Fig. 5. These contours are based on an independent dataset of thickness measurements (different from the assimilation dataset). The presence of the distal secondary thickness maximum was reproduced by the reconstructed deposits taking into account the ash aggregation processes in the numerical simulation. To this purpose, the parameterisation proposed (Cornell et al., 1983; Folch et al., 2010) was used assuming an aggregate particle class having a diameter of 200 μm and a density sampled in a range centred around 450 $\mathrm{kg\,m^{-3}}$ (see Table 4).
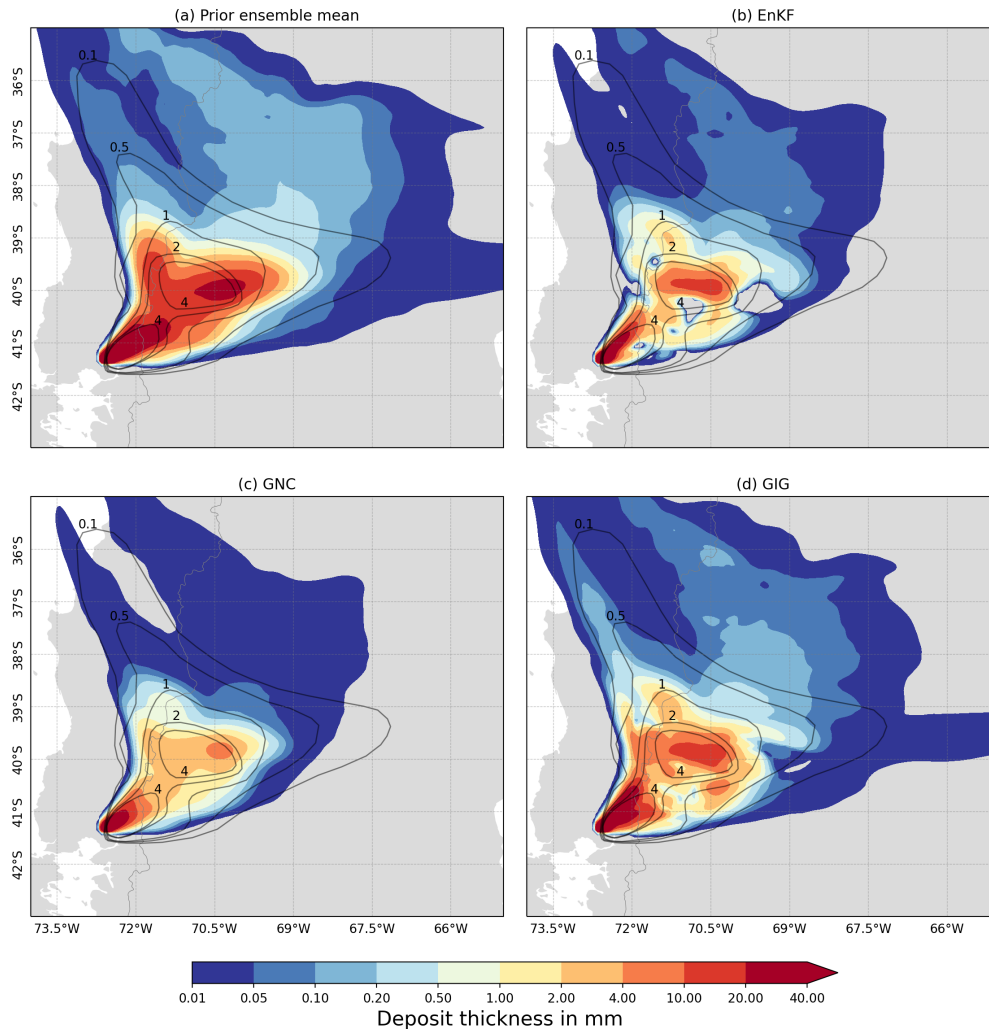
Figure 5. Reconstructed tephra fall deposit according to the Forecast (a), EnKF method (b), GNC method (c) and GIG method (d). The Romero et al. (2016) deposit contours (isopachs for 0.1, 0.5, 1, 2 and 4 mm) are also over imposed for comparative purposes.

420   ~~Figure ?? compares the analysis results at each sampling site with observations (from the same assimilated dataset)according to the GNC (~~

The forecast (Fig. ~~??a) and the GIG~~ 5a) indicates an excessively high deposited mass. In particular, the ~~of~~ 2- and 4-mm contours are overestimated by almost an order of magnitude. This is corrected by the three methods (Fig. ~~??b)methods. Given that the GIG method also provides the mean and standard deviation of~~ 5b-c), which yield analyses with reduced total mass on

425   the ground.

The EnKF analysis shows a very noisy spatial distribution with large oscillations and negative values in some regions leading to artificial spatial structures. On the other hand, the GNC shows smoother deposit thickness contours with a spatial distribution

having a more physically-plausible structure. The GIG method represents an intermediate case between the EnKF and GNC methods. Although this method gives unrealistic results as well, the fraction of negative values and the amplitude of oscillations

430 are noticeably reduced compared to the ~~analysis, this information is used in Fig. ??b to represent the analysis error through error bars. Unfortunately, an estimate~~ EnKF method (negative data was remove and reassigned to zero).

3.6 Validation

In order to evaluate the performance of the analysis ~~error is not provided by the GNC method and , consequently, error bars cannot be shown in Fig. ??a. Note that most of the data lies in the 1:10 ratio band, with only 4 (2) data points laying outside~~

435 ~~this band for the GNC (GIG) method despite the challenging fact that~~ schemes, the full dataset of observations was split into two subsets: an assimilation dataset and a validation dataset. The assimilation dataset was used to produce new analyses and the validation metrics defined in Sect. 3.2 were computed on the validation dataset. This methodology ensures that validation is done against non-assimilated observations. However, the validation metrics will be meaningless if the assimilation and validation datasets are strongly correlated (e.g.if measurement sites are very close to each other). The splitting procedure aims

440 to reduce the correlation between both subsets and is described in the supplementary material.

Figure 6 compares the analysis results at each sampling site with observations from the assimilation dataset (left column panel) and the validation dataset (right column panel) considering a dataset partition of 60% and 40%, respectively. The forecast systematically overestimates observations (Fig. 6a,e), as already noted in Sect. 3.5. Since measurements span a range of four orders of magnitude between 10^{-2} and 10^{2} ~~em, approximately.~~ cm, a good agreement for the entire range of data values turns to

445 be challenging. Nevertheless, results for the analyses corresponding to EnKF (Fig. 6b,f), GNC (Fig. 6c,g), and GIG (Fig. 6d,h) methods show that most of the data lies within the 1-to-10 ratio band (dashed black line).

~~Table ?? reports the different evaluation metries in~~ In order to quantify deviations from observations ~~in~~ according to the prior ensemble mean (forecast) and the ~~analyses for both GNC and GIG methods. The bias computed using~~ analysis schemes, the evaluation metrics computed on the assimilation dataset ~~for the prior mean was negative (MBE=-4.31 cm) , meaning~~

450 ~~that measurements were overestimated systematically by the first guess. In contrast, the analysis underestimates observations according to GNC method (MBE=1.32 cm) and slightly overestimates according to GIG methos (MBE=-0.30 cm). In both cases, the magnitude of the forecast bias was clearly reduced. The prior RMSE (RMSE=8.53 cm) was also reduced by the GNC (RMSE=4.23 cm)~~ 60%) and the validation dataset (40%) are reported in Fig. 7. Since the GIG method is stochastic, six realisations of the analysis were computed and the average over the realisations are reported for each metric. The EnKF,

455 GNC and ~~the GIG (~~GIG methods improve all metrics compared to the forecast. In particular, the RMSE ~~=2.02 cm)methods. The impact of the assimilation is better characterised by the relative or weighted metrics.The assimilation methods lead to astrong reduction of both wMBE and wRMSE, as reported in Table ??. For example, the weighted RMSE was reduced from wRMSE=3.96 to wRMSE=1.15 (GNC method) and wRMSE=1.31 (GIG method~~computed from the assimilation dataset (Fig. 7a, green bar) is reduced from 7.1 (forecast) to 0.9 (EnKF) and 1.2 (GNC,GIG). Since the EnKF provides the best estimate

460 in terms of RMSE, this is an expected result. However, when the RMSE is computed on the validation dataset (Fig. 7a, brown bar), the performance of EnKF (2.3) and GIG (5.2) methods degrades significantly and the best performance is achieved by
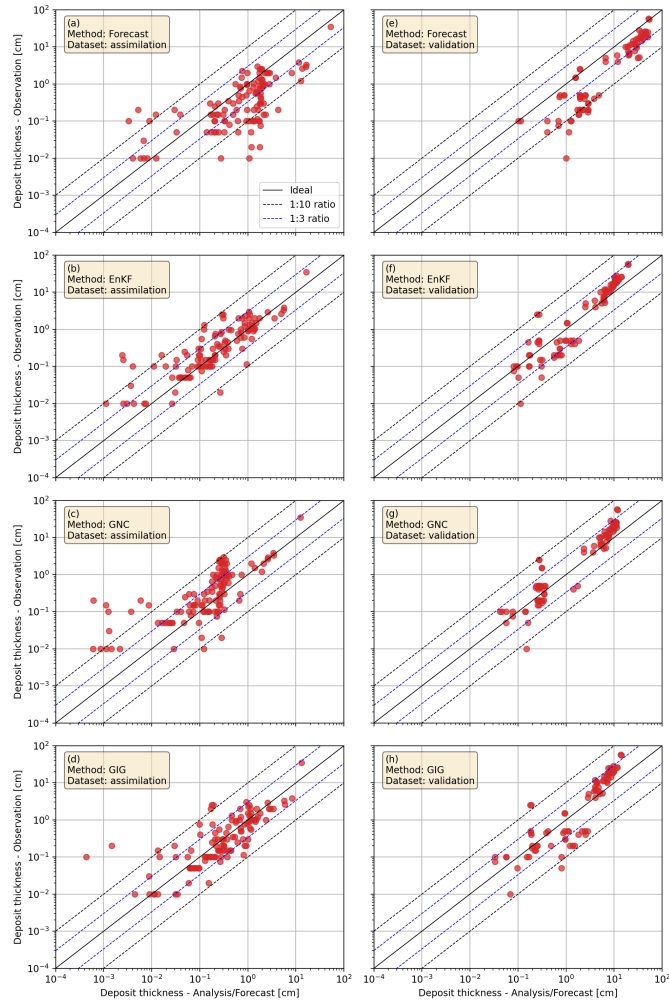
Figure 6. Comparison between ~~analysis~~ analyses and observations ~~from~~ for the forecast (a-b), EnKF (c-d), GNC (e-f) and GIG (g-h) methods. Panel on the left column present the assimilation dataset (~~161 points, see Table 2~~60% of the full dataset) and the validation dataset (40% of the full dataset) is shown on the right column.

the GNC method (1.1). Note that ~~the analysis are~~ this result is very close to the ideal value ~~wRMSE~~RMSE=~~1.~~1, meaning that deviations are similar to the observation uncertainty (see Eq. 24b).

~~Metrics computed for each method considering the assimilation and validations datasets. The parameter ρ_o used for the GIG~~

465 ~~method is specified in parentheses (see Eq. ??).~~ The bias is presented in Fig. 7b. The GNC MBE (0.3) is positive, meaning that measurements are underestimated, but this bias is within the observation uncertainty range since MBE<1. In contrast, the analysis overestimates observations according to EnKF (MBE~~RMSEwMBE wRMSE (cm)(cm) Prior mean -4.31 8.53 -2.15 3.96 GNC Method 1.32 4.23 0.30 1.15 GIG Method† -0.30 2.02 -0.18 1.31 Prior mean -2.11 5.39 -3.77 6.47 GNC Method~~

1.43 6.21 -0.04 1.81 GIG Method (0.00) † 0.11 2.51 -0.55 1.83 GIG Method (0.01)† 0.22 2.97 -0.64 2.05 GIG Method (0.05)‡ 0.23 3.28 -0.60 1.96 GIG Method (0.10)† 0.00 1.90 -0.65 1.94 GIG Method (0.50)† -0.07 2.73 -0.68 2.34 GIG Method (0.00)§ 0.17 2.85 -0.72 2.32 GIG Method (0.01)§ 0.21 2.98 -0.68 2.36 GIG Method (0.05)§ 0.25 3.23 -0.73 2.52 GIG Method (0.10)§ 0.11 2.51 -0.78 2.44 GIG Method (0.50)§ 0.05 2.83 -0.77 2.55

For the validation dataset, the metrics for the prior mean yield values similar to those of the assimilation dataset, wMBE=-3.77 and wRMSE-0.5) and GIG (MBE=6.47. In terms of the weighted metrics, the GNC method showed the best performance (wMBE-1.5) methods. The best SMAPE results were obtained for GNC (SMAPE=-0.04 and wRMSE31.8%) and EnKF (SMAPE=1.81) 32.6%) as shown in Fig. 7c. Finally, Fig. 7d presents the 1-to-3 Ratio Band Score, defined as the percentage of data points within the 1-to-3 ratio band (dashed blue line in Fig. 6). Regarding the GIG method, the metrics were computed for different values of the parameter ρ_o, required by (??) Again, $\rho_o = 0.0$ (no localisation) , $\rho_o = 0.01$, $\rho_o = 0.05$, $\rho_o = 0.1$ and $\rho_o = 0.5$. Since the GIG method is stochastic, five realisations were performed for each ρ_o. Table ?? reports the metrics corresponding to the best realisation (the one with the least wRMSE)and the average over the realisations. Taking into account both weighted bias and RMSE, the best performance result was obtained for $\rho_o = 0.0$, wMBE=-0.55 and wRMSE=1.83, suggesting that the localisation procedure did not have a significant impact. This is also confirmed by looking at the average wRMSE: the best result is obtained when localisation is disabled (wRMSE=2.32 for $\rho_o = 0$the GNC method (84.1%).

Figure ?? compares the analyses with measurements from the

In conclusion, the GNC method outperforms the EnKF and GIG methods in term of all metrics computed on the validation dataset. While most of the points lie within the 1:10 ratio band, an almost zero thickness was found according to the analysis at four sampling sites with positive measurements. These points were excluded from the log-log plot in Fig. ?? in order to facilitate the visualisation, but they were included in the calculation of the evaluation metrics reported in Table ??. The sampling sites for the excluded data are found west of the Andes as indicated by In contrast, the EnKF and GIG analyses perform well over regions around the observation sites, but the analyses cannot fully capture the entire deposit features beyond these regions. In order to illustrate this point, the uncorrected EnKF analysis (i.e.negative data was not removed) and the location of the assimilated observations (60% of the full dataset) is presented in Fig. 2 (cross symbols), suggesting that the tephra deposit distribution was not properly modelled over this area with strong gradients of mass loading8, where colour shaded regions represent positive data and negative data are masked. Large oscillations emerge in regions without observation data leading to artificial structures in the deposit field. This example clearly highlights the problems arising when the EnKF is applied to this case and the importance of dealing with the assimilation of volcanic deposit data using alternative approaches.

Finally, the tephra fall deposit reconstruction according to GNC/GIG methods is To conclude this section, different partitions of the observational dataset are considered and the RMSE is computed on the validation dataset. Results are shown in Fig. 5. For comparative purposes, 9, expressed in terms of the percentage of assimilated observations for each partition. Again, the results show that the GNC outperforms the EnKF and GIG methods systematically, not just for a particular choice of the validation dataset. Unfortunately, it is not possible to conclude that the Romero et al. (2016) deposit contours (hand-drawn isopachs for 0.1, 0.5, 1.0 and 2.0 mm) are over imposed in Fig. 5, providing an alternative estimate of the fallout deposit from an independent datasetof thickness measurements. Note that a remarkable feature of the GNC/GIG reconstructed deposits is
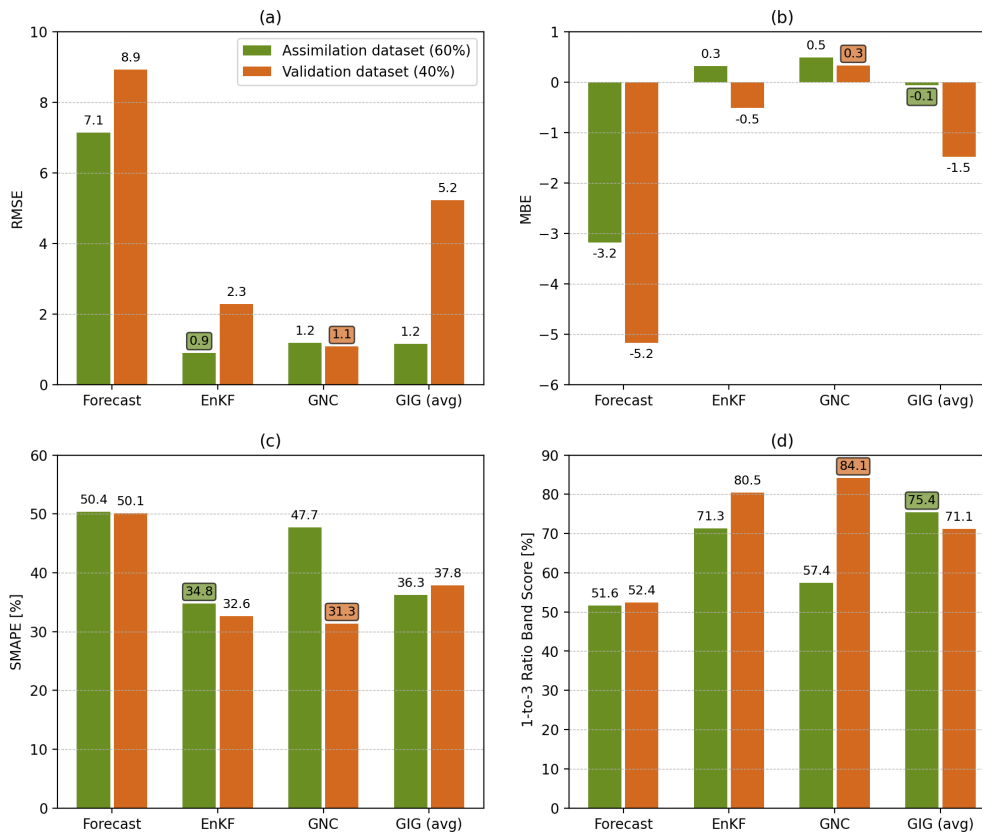
Figure 7. Evaluation metrics for the prior ensemble mean (forecast) and the EnKF, GNC and GIG analysis schemes. The metrics were computed on the assimilation and validation datasets considering a partition of the full dataset of 60% and 40%, respectively.

505 ~~the presence of a distal secondary thickness maximum ~250 km downwind from vent, indicating that this volcanic eruption involved a complex plume dynamics. Both reconstructions successfully captured the location of this maximum due to ash aggregation processes which were described using the parameterisation proposed (Cornell et al., 1983; Folch et al., 2010) assuming a aggregate particle class having a diameter of 200 and a density sampled in a range centred around 450 (see Table 4)~~GIG method has improved the results of the EnKF analysis from the evidence presented in this paper.

4   Source term inversion: Application to the 2015 Calbuco eruption

510 A major advantage of the GNC method is that it allows estimating the Eruption Source Parameters (ESP) in a straightforward way, with inverse modelling coming at no extra computational cost. This is because FALL3D solves an almost linear problem with weak non-linearity effects (e.g. due to gravity current, wet deposition, or aggregation) and, consequently, a re-scaling of the emission source term s_i associated with the i-th ensemble member according to $s_i \rightarrow w_i s_i$ leads to a deposit mass loading
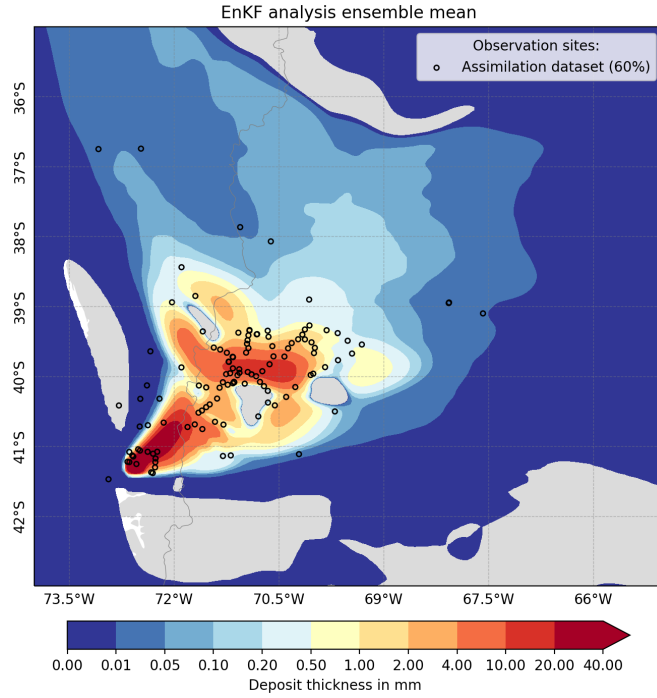
Figure 8. ~~Comparison between analysis~~ Tephra fall deposit according to the EnKF method and location of the assimilated observations ~~from~~ (60% of the ~~validation~~ full dataset~~(41 points, see Table 2 and text for explanation)~~. Colour shaded regions represent positive data and negative data is masked in order to illustrate the emergence of large oscillations and unphysical values in regions with scarce observational data.

re-scaled correspondingly, being w_i the weight factor provided by the GNC method. In this case, the best estimation of the

515   total source term is given by:

$$s^a = \sum_i w_i s_i \tag{27}$$

where $\{s_i\}$ are the emission source terms of the prior ensemble members (in $\mathrm{kg\,m^{-3}\,s^{-1}}$). Figure 10 shows the emission rate profiles resulting from the source inversion, expressed in terms of the linear source emission strength (in $\mathrm{kg\,m^{-1}\,s^{-1}}$), i.e. $s^a \times dA$, where dA is the area of the cell grid.

520   As most of the 256 weight factors converge to zero, $w_i \to 0$, the profiles in Fig. 10 reflect only those ensemble members that effectively contribute to the analysed deposit mass loading. According to the GNC inverse modelling results, each eruptive phase is characterised by different vertical mass distribution and emission rates. The first phase results on higher cloud top heights, reaching almost 20 km (a.s.l.) whereas the column heights during the second phase remain at around 16 km (a.s.l.). Note also that the prior ensemble was defined assuming the same emission source and sampling parameters for both eruptive

525   phases. For instance, the ensemble configuration was defined assuming the same reference value for column heights, i.e. 15 km above vent level for both phases (Table 4) or 17 km (a.s.l.), approximately. Therefore, the resulting asymmetry between both
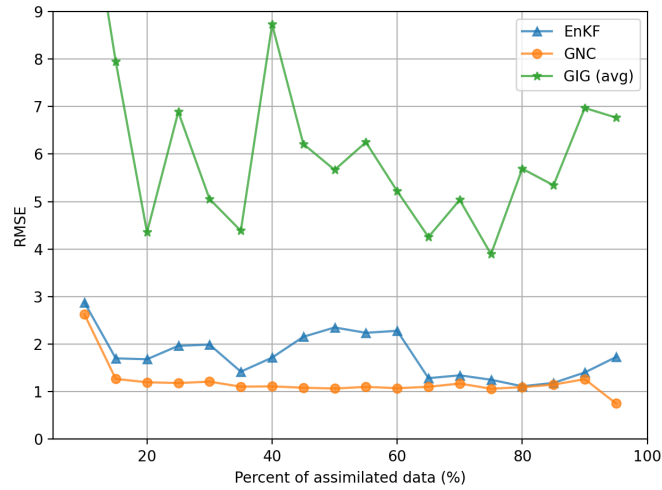
Figure 9. ~~Reconstructed tephra fall deposit according to GNC (a) and GIG (b) methods. The Romero et al. (2016) deposit contours (isopachs~~ RMSE computed on the validation dataset for ~~0.1, 0.5, 1.0 and 2.0 mm) are also over imposed for comparative purposes~~different partitions of the full dataset of observations expressed in terms of the percentage of assimilated observations.
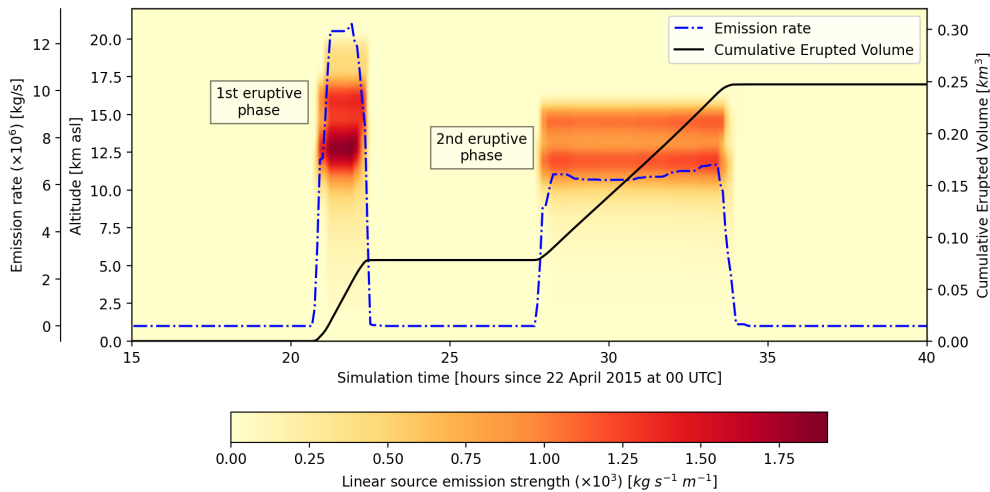
Figure 10. Profiles of emission rate and time-series of Eruption Source Parameters (ESP) for the 2015 Calbuco eruption according to the GNC inverse modelling approach. The solid line represents the cumulative erupted volume in km^3 and the dash-dotted line indicates the mass emission rate $(\times 10^6)$ in $kg\,s^{-1}$.

pulses observed in Fig. 10 is a ~~result~~ consequence of the GNC inversion procedure. In other words, the GNC method can discard inappropriate ensemble members and pick out those that are consistent with observations. ~~The vertical mass distribution for~~ ~~the first phase results on a larger spread and~~

The time-series for mass eruption rate and total erupted volume are also depicted in Fig. 10. Although the maximum emission rate is reached during the more intense first phase, most of the total erupted mass stems from the longer second phase. ~~Assuming~~ In order to estimate the total erupted volume, a bulk density of $\rho_b =800\ \mathrm{kg\,m^{-3}}$ ~~, it is possible to estimate the total erupted volume. The time-series for mass eruption rate and total erupted volume are also depicted in Fig. 10.~~ was assumed (a typical value for fresh deposits). In particular, the final total erupted volume was around ~~0.33~~0.25 $\mathrm{km^3}$. According to the inversion, the erupted volume corresponding to the first and second phases were ~~0.11~~0.078 $\mathrm{km^3}$ (~~32.9~~31.6%) and ~~0.22~~0.169 $\mathrm{km^3}$ (~~67.1~~68.4%), respectively. These results are in good agreement with the estimations reported by Romero et al. (2016), which give a total bulk tephra fall deposit volume of $0.27 \pm 0.007\ \mathrm{km^3}$ with 62% of the total volume corresponding to the second phase.

5 Discussion

Traditional ensemble-based DA methods such as the ensemble Kalman filter (EnKF) are based on the Gaussian hypothesis. However, it is well known that analyses produced by these methods are suboptimal when either the model state variables or the observation errors are not Gaussian distributed. Volcanic aerosol concentrations and tephra deposit mass loading are two remarkable examples of non-Gaussian-distributed variables with highly-skewed distributions (Mingari et al., 2022). This explains why the application of EnKF-like methods in VATD models often leads to nonphysical results and oscillations (e.g. the occurrence of negative concentrations). The ensemble-based GNC ~~and GIG methods~~ method introduced in this paper ~~are promising~~ and the GIG method (Bishop, 2016) represent potential alternatives for dealing with the assimilation of volcanic data. ~~These~~ In addition, the proposed methodologies could be beneficial beyond volcanic tephra, for example in situations involving non-negative variables with right-skewed probability distributions, such as water-vapour mixing ratio, rainfall or aerosol concentrations. The GNC and GIG methods differ in the assumptions made about the prior distribution and the likelihood of the observation conditioned on the true state. Both methods have been applied to the assimilation of volcanic deposit data and ~~have shown to have a similar performance in terms of the~~ the results were compared to the traditional EnKF analysis by means of different evaluation metrics. The GNC method produced physically plausible results and significantly outperformed the other methods and the prior ensemble mean. Unfortunately, it was not possible to conclude that the GIG method improved the EnKF analysis from the results presented here.

The GNC method assumes a multi-dimensional Gaussian distribution and solves an optimization problem with nonnegative constraints to ensure plausible physical solutions. The GNC method, constrained here to assimilate deposit observations, can be easily extended to other observables as long as the observation operator is linear. For example, VATD models could use it to assimilate column mass observations of volcanic aerosols, but the assimilation of other satellite-retrieved variables (e.g. Aerosol Optical Depth) would require of an alternative approach. ~~Since the linear estimator (11) includes the prior ensemble mean ($w_i = 1/m$) and any specific ensemble member ($\boldsymbol{w} = [0,\ldots,1,\ldots,0]^\intercal$) among the possible solutions, the~~ The solution obtained through the minimisation process (17) converges to an analysis state which, by construction, improves the prior ensemble mean and any individual ensemble member~~.~~ since the linear estimator (11) includes, among the possible solutions, the prior ensemble

mean ($w_i = 1/m$) and any specific ensemble member (e.g.$w_i = \delta_{ij}$ is solution for the j-th member). Furthermore, the GNC method ensures that the prior ~~wRMSE~~ RMSE is reduced by the analysis state. This can be checked from (10) by noting that ~~, if~~ the weighted RMSE of the prior ensemble mean is the initial value of the normalised cost function (defined as $\sqrt{J/p}$), provided that the iterative solving procedure is started from the uniform vector with components $w_i = 1/m$ ~~, the normalised cost function (defined as $\sqrt{J/p}$) coincides with the weighted RMSE of the prior ensemble mean before the first iteration (as long as the~~ and the matrix $\mathbf{R}$ is diagonal~~). In fact, the metrics reported in Table ?? illustrate~~. Figure 7a illustrates this property of the solution, e.g. the ~~wRMSE is reduced from wRMSE~~RMSE computed on the validation dataset was reduced from RMSE=~~8.53~~8.9 (prior ensemble mean) to ~~wRMSE~~RMSE=~~1.32 (analysis). This is not the case for individual ensemble members; the weighted RMSE associated with a specific member may~~1.1 (GNC). In contrast, the analysis RMSE is not ensured to be less than the ~~analysis relative error~~RMSE associated with individual ensemble members. This is a desirable property of the solution for statistically non-significant members. In fact, note that the first term in (10) penalises deviations from the ensemble mean (i.e. statistically non-significant members), while the second term penalises deviations from the observations. As a result, the solution provided by the GNC method satisfies two properties: (i) the analysis is statistically significant and, (ii) deviations from the observations are small.

The GIG method is a sequential assimilation procedure proposed by Bishop (2016), in which single observations are sequentially assimilated. The GIG method is based on the GIG equation set for the special case where the prior distribution can be described by a gamma PDF and ~~an inverse-gamma observation likelihood~~ the observation likelihood by an inverse-gamma PDF. This is a stochastic method providing an ensemble of analyses and does not require a linear observation operator, in principle. These reasons make the GIG method a ~~better~~ good candidate for implementation in VATD models as it would allow performing multiple assimilation cycles by restarting ~~a corrected ensemble forecast~~the model from the analysis ensembles. The GIG method enables near-zero semi-positive-definite variables with highly skewed uncertainty distributions to be assimilated, and avoids the occurrence of negative mass loading at the observation site. ~~However, a linear regression approximation is used to propagate~~ However, the observation information is propagated to the extended model-observation state vector ~~in the assimilation step. Consequently, negative values can be introduced by~~ using the linear regression approximation (22)~~when the analysis in computed at the grid points of the computational domain. In order to solve this potential issue, we explored a localisation technique which limits the observation impacts based on their background error correlations using (??). Unfortunately, this localisation approach did not show a clear improvement according to the evaluation metrics . Probably.~~ This approach introduces artificial structures in the spatial distribution of the deposit beyond the observation sites, including a few negative values, and the validation metrics degrades significantly over regions with scarce observational data assimilated. As a conclusion, the linear regression approximation ~~went well in this work due to the linearity of the observation operator. However, for more general problems involving non-linear observation operators, the regression (22) may need to be reformulated and the localisation approach proposed here could be more relevant.~~ should be reformulated or corrected, e.g.exploring localisation techniques, in order to enhance the quality of the deposit reconstruction using the GIG method.

6 Conclusions

This paper has proposed two ensemble-based data assimilation methods for semi-positive-definite variables. The methods were applied to reconstruct the tephra fallout deposit of the 2015 Calbuco eruption in Chile by assimilating measurements of deposit thickness. An assessment based on an independent ~~observational dataset yielded similar results for both~~ validation dataset was carried out for the GNC/GIG methods in terms of ~~the evaluation metrics . We conclude that both methods are~~ different evaluation metrics and the results compared to two references: the ensemble prior mean and the EnKF analysis.

The evidence from this study suggests that the GNC method was the most skilful approach and represent a promising alternative for assimilation of volcanic fallout data. ~~In addition, the techniques presented in this work can be extended to other volcanic observations, including satellite-retrieved data of different aerosol types, and used in operational forecast contexts.~~

The GNC method provides an ensemble of weight factors and can be used also for source term inversion in a straightforward way. Unlike the majority of source term inversion methods ~~(e.g. ?)~~(e.g. Folch and Mingari, 2023), which focus on determining specific ESP associated with oversimplified parameterisations of the source term, this approach reconstructs the overall space-time distribution of the source and it is not constrained by any specific parameterisation of the emission source term.

On the other hand, although an interesting approach, the GIG method failed to improve the EnKF analysis. Evidently, the linear regression used by the GIG method needs to be reformulated or corrected. The GIG method is a second-order method and provides an ensemble of analyses without the linear observation operator assumption. Consequently, it represents an attractive alternative for assimilation of volcanic aerosol observations from satellite retrievals. To this purpose, the analysis ensemble from the GIG method could be used to perform multiple assimilation cycles by restarting an ensemble forecast. This approach has the potential to improve the accuracy of operational forecasts of volcanic clouds. In its present form, the GNC method is not suitable for data assimilation of volcanic aerosol observations in a context of operational forecasting as it does not provide an analysis ensemble. To achieve this, future studies should focus on extending the method in order to formulate a second-order analysis scheme based on the GNC method.

7 ~~Generation of the analysis ensemble in the GIG method~~

~~In order to generate an ensemble y_{ji}^a with the correct posterior density from the forecast ensemble for the case of an univariate gamma prior and an inverse-gamma observation-likelihood, Bishop (2016) proposed the following stochastic equation:~~

$$\frac{y_{ji}^a - \overline{y}_j^a}{\overline{y}_j^a} = \frac{y_{ji}^f - \overline{y}_j^f}{\sqrt{\left(\overline{y}_j^f\right)^2 + \mathrm{var}(y_{ji}^f)}} + \tilde{P}_j^r \left(\tilde{P}_j^r + \tilde{R}_j^r\right)^{-1} \times$$

$$\times \left\{ \frac{z_{ji} - \overline{z}_j}{\sqrt{\left(\overline{z}_j\right)^2 - 2\mathrm{var}(z_{ji})}} - \frac{y_{ji}^f - \overline{y}_j^f}{\sqrt{\left(\overline{y}_j^f\right)^2 + \mathrm{var}(y_{ji}^f)}} \right\}$$

where $\overline{y}_j^a$ can be computed using Eq. (19a). Equation (20) ensures that the analysis ensemble y_{ji}^a is consistent with the type 1 relative error variance of the posterior given by Eq. (19b) provided z_{ji} is randomly sampled from a gamma PDF with type 1 relative error variance R_j^z and mean $\overline{z}_j$ given by:

$$\left(R_j^z\right)^{-1} = \left(\tilde{R}_j^r\right)^{-1} + 2$$

$$\overline{z}_j = \frac{\left(\tilde{R}_j^r\right)^{-1} + 2}{\left(\tilde{R}_j^r\right)^{-1}} y_j^o$$

Code and data availability. FALL3D-8.2.0 is available under the version 3 of the GNU General Public License (GPL) at https://doi.org/10.5281/zenodo.6343786 (Folch et al., 2020; Prata et al., 2021). Observational datasets, code used for the GNC/GIG methods and input model parameter file along with the pre- and post-processing scripts have been archived on Zenodo at https://doi.org/10.5281/zenodo.7259531.

Author contributions. Conceptualisation, L.M., A.C.; Methodology, L.M.; Software, L.M., A.F., G.M.; Resources, L.M.; Writing—original draft, L.M.; Writing—review and editing, L.M., A.C., A.F., G.M.; Visualisation, L.M.; Supervision, A.C., A.F.; Funding Acquisition, A.F. All authors have read and approved the final version of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work has been partially funded by the H2020 Center of Excellence for Exascale in Solid Earth (ChEESE) under the Grant Agreement No. 823844. 823844 and by the European Union's Horizon Europe Research and Innovation Program under grant agreement No. 101058129 (DT-GEO). We acknowledge the Partnership for Advanced Computing in Europe (PRACE) for awarding us access to the Joliot-Curie supercomputer at the CEA's Very Large Computing Center (TGCC, France). We thank Alexa Van Eaton from USGS for providing us with the assimilation dataset and the digitalised isopach contours. We are also grateful to Florencia Reckziegel for kindly providing us with the validation her dataset.

References

Amezcua, J. and Leeuwen, P. J. V.: Gaussian anamorphosis in the analysis step of the EnKF: a joint state-variable/observation approach, Tellus A: Dynamic Meteorology and Oceanography, 66, 23 493, https://doi.org/10.3402/tellusa.v66.23493, 2014.

Anderson, J. L.: A Local Least Squares Framework for Ensemble Filtering, Mon. Weather Rev., 131, 634–642, https://doi.org/10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2, 2003.

Beckett, F. M., Witham, C. S., Leadbetter, S. J., Crocker, R., Webster, H. N., Hort, M. C., Jones, A. R., Devenish, B. J., and Thomson, D. J.: Atmospheric Dispersion Modelling at the London VAAC: A Review of Developments since the 2010 Eyjafjallajökull Volcano Ash Cloud, Atmosphere, 11, https://doi.org/10.3390/atmos11040352, 2020.

Bishop, C. H.: The GIGG-EnKF: ensemble Kalman filtering for highly skewed non-negative uncertainty distributions, Q. J. Roy. Meteor. Soc., 142, 1395–1412, https://doi.org/10.1002/qj.2742, 2016.

Bonadonna, C. and Costa, A.: Plume height, volume, and classification of volcanic explosive eruptions based on the Weibull function, Bull. Volcanol., 75, 1–19, https://doi.org/10.1007/s00445-013-0742-1, 2013.

Bonadonna, C., Ernst, G., and Sparks, R.: Thickness variations and volume estimates of tephra fall deposits: the importance of particle Reynolds number, J. Volcanol. Geoth. Res., 81, 173–187, https://doi.org/10.1016/S0377-0273(98)00007-9, 1998.

Bonadonna, C., Folch, A., Loughlin, S., and Puempel, H.: Future developments in modelling and monitoring of volcanic ash clouds: outcomes from the first IAVCEI-WMO workshop on Ash Dispersal Forecast and Civil Aviation, Bull. Volcanol., 74, 1–10, 2012.

Bonadonna, C., Biass, S., and Costa, A.: Physical characterization of explosive volcanic eruptions based on tephra deposits: Propagation of uncertainties and sensitivity analysis, J. Volcanol. Geoth. Res., 296, 80–100, https://doi.org/10.1016/j.jvolgeores.2015.03.009, 2015.

Burgers, G., Van Leeuwen, P. J., and Evensen, G.: Analysis scheme in the ensemble Kalman filter, Mon. Weather Rev., 126, 1719–1724, 1998.

Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G.: Data assimilation in the geosciences: An overview of methods, issues, and perspectives, WIREs Clim. Change, 9, e535, https://doi.org/10.1002/wcc.535, 2018.

Chang, C.-C. and Kalnay, E.: Applying prior correlations for ensemble-based spatial localization, Nonlinear Proc. Geoph., 29, 317–327, https://doi.org/10.5194/npg-29-317-2022, 2022.

Clarkson, R. J., Majewicz, E. J., and Mack, P.: A re-evaluation of the 2010 quantitative understanding of the effects volcanic ash has on gas turbine engines, P. I. Mech. Eng. G-J. Aer., 230, 2274–2291, https://doi.org/10.1177/0954410015623372, 2016.

Constantinescu, R., White, J. T., Connor, C. B., Hopulele-Gligor, A., Charbonnier, S., Thouret, J.-C., Lindsay, J. M., and Bertin, D.: Uncertainty Quantification of Eruption Source Parameters Estimated From Tephra Fall Deposits, Geophys. Res. Lett., 49, e2021GL097 425, https://doi.org/10.1029/2021GL097425, 2022.

Cornell, W., Carey, S., and Sigurdsson, H.: Computer simulation and transport of the Campanian Y-5 ash, J. Volcanol. Geoth. Res., 17, 89–109, https://doi.org/10.1016/0377-0273(83)90063-X, 1983.

Costa, A., Folch, A., and Macedonio, G.: A model for wet aggregation of ash particles in volcanic plumes and clouds: I. Theoretical formulation, J. Geophys. Res.-Sol. Ea., 115, https://doi.org/10.1029/2009JB007175, 2010.

Costa, A., Pioli, L., and Bonadonna, C.: Assessing tephra total grain-size distribution: Insights from field data analysis, Earth Planet. Sc. Lett., 443, 90–107, https://doi.org/10.1016/j.epsl.2016.02.040, 2016.

Dacre, H. F. and Harvey, N. J.: Characterizing the Atmospheric Conditions Leading to Large Error Growth in Volcanic Ash Cloud Forecasts, J. Appl. Meteorol. Climatol., 57, 1011–1019, https://doi.org/10.1175/JAMC-D-17-0298.1, 2018.

Dare, R. A., Smith, D. H., and Naughton, M. J.: Ensemble Prediction of the Dispersion of Volcanic Ash from the 13 February 2014 Eruption of Kelut, Indonesia, J. Appl. Meteorol. Climatol., 55, 61–78, https://doi.org/10.1175/JAMC-D-15-0079.1, 2016.

Degruyter, W. and Bonadonna, C.: Improving on mass flow rate estimates of volcanic eruptions, Geophys. Res. Lett., 39, https://doi.org/10.1029/2012GL052566, 2012.

Denlinger, R. P., Pavolonis, M., and Sieglaff, J.: A robust method to forecast volcanic ash clouds, J. Geophys. Res.-Atmos., 117, https://doi.org/10.1029/2012JD017732, 2012.

Dominguez, L., Bonadonna, C., Forte, P., Jarvis, P. A., Cioni, R., Mingari, L., Bran, D., and Panebianco, J. E.: Aeolian remobilisation of the 2011-Cordón Caulle Tephra-Fallout Deposit: example of an important process in the life cycle of Volcanic Ash, Front. Earth Sci., 7, 343, 2020.

Durant, A. J., Rose, W. I., Sarna-Wojcicki, A. M., Carey, S., and Volentik, A. C. M.: Hydrometeor-enhanced tephra sedimentation: Constraints from the 18 May 1980 eruption of Mount St. Helens, J. Geophys. Res.-Sol. Ea., 114, https://doi.org/10.1029/2008JB005756, 2009.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res.-Oceans, 99, 10 143–10 162, https://doi.org/10.1029/94JC00572, 1994.

Folch, A. and Mingari, L.: Data assimilation of volcanic clouds: recent advances and implications on operational forecasts, in: Applications of Data Assimilation and Inverse Problems in the Earth Sciences, edited by Ismail-Zadeh, A., Castelli, F., Jones, D., and Sanchez, S., Cambridge University Press, Cambridge, 2023.

Folch, A., Costa, A., Durant, A., and Macedonio, G.: A model for wet aggregation of ash particles in volcanic plumes and clouds: II. Model application, J. Geophys. Res.-Sol. Ea., 115, https://doi.org/10.1029/2009JB007176, 2010.

Folch, A., Mingari, L., Osores, M. S., and Collini, E.: Modeling volcanic ash resuspension - application to the 14–18 October 2011 outbreak episode in central Patagonia, Argentina, Nat. Hazards Earth Syst. Sci., 14, 119–133, https://doi.org/10.5194/nhess-14-119-2014, 2014.

Folch, A., Mingari, L., Gutierrez, N., Hanzich, M., Macedonio, G., and Costa, A.: FALL3D-8.0: a computational model for atmospheric transport and deposition of particles, aerosols and radionuclides – Part 1: Model physics and numerics, Geosci. Model Dev., 13, 1431–1458, https://doi.org/10.5194/gmd-13-1431-2020, 2020.

Folch, A., Mingari, L., and Prata, A. T.: Ensemble-Based Forecast of Volcanic Clouds Using FALL3D-8.1, Front. Earth Sci., 9, https://doi.org/10.3389/feart.2021.741841, 2022.

Folch, A., Abril, C., Afanasiev, M., Amati, G., Bader, M., Badia, R. M., Bayraktar, H. B., Barsotti, S., Basili, R., Bernardi, F., et al.: The EU Center of Excellence for Exascale in Solid Earth (ChEESE): Implementation, results, and roadmap for the second phase, Future Gener. Comp. Sy., pp. 47–61, https://doi.org/j.future.2023.04.006, 2023.

Fu, G., Heemink, A., Lu, S., Segers, A., Weber, K., and Lin, H.-X.: Model-based aviation advice on distal volcanic ash clouds by assimilating aircraft in situ measurements, Atm. Chem. Phys., 16, 9189–9200, https://doi.org/10.5194/acp-16-9189-2016, 2016.

Fu, G., Prata, F., Lin, H. X., Heemink, A., Segers, A., and Lu, S.: Data assimilation for volcanic ash plumes using a satellite observational operator: a case study on the 2010 Eyjafjallajökull volcanic eruption, Atm. Chem. Phys., 17, 1187–1205, https://doi.org/10.5194/acp-17-1187-2017, 2017.

Harvey, N. J., Dacre, H. F., Webster, H. N., Taylor, I. A., Khanal, S., Grainger, R. G., and Cooke, M. C.: The impact of ensemble meteorology on inverse modeling estimates of volcano emissions and ash dispersion forecasts: Grímsvötn 2011, Atmosphere, 11, 1022, https://doi.org/10.3390/atmos11101022, 2020.

Husak, G. J., Michaelsen, J., and Funk, C.: Use of the gamma distribution to represent monthly rainfall in Africa for drought monitoring applications, Int. J. Climatol., 27, 935–944, https://doi.org/10.1002/joc.1441, 2007.

715 Jazwinski, A. H.: Stochastic processes and filtering theory, Academic Press, Inc., New York, 1970.

Kalman, R. E.: A New Approach to Linear Filtering and Prediction Problems, J. Basic Eng.-T. ASME, 82, 35–45, https://doi.org/10.1115/1.3662552, 1960.

Kliewer, A. J., Fletcher, S. J., Jones, A. S., and Forsythe, J. M.: Comparison of Gaussian, logarithmic transform and mixed Gaussian–log-normal distribution based 1DVAR microwave temperature–water-vapour mixing ratio retrievals, Q. J. Roy. Meteor. Soc., 142, 274–286,
720 https://doi.org/10.1002/qj.2651, 2016.

Madankan, R., Pouget, S., Singla, P., Bursik, M., Dehn, J., Jones, M., Patra, A., Pavolonis, M., Pitman, E., Singh, T., and Webley, P.: Computation of probabilistic hazard maps and source parameter estimation for volcanic ash transport and dispersion, J. Comput. Phys., 271, 39–59, https://doi.org/10.1016/j.jcp.2013.11.032, 2014.

Martí, A., Folch, A., Costa, A., and Engwell, S.: Reconstructing the plinian and co-ignimbrite sources of large volcanic eruptions: a novel
725 approach for the Campanian Ignimbrite, Sci. Rep., 6, 1–11, https://doi.org/10.1038/srep21220, 2016.

Martinez, B. M., Luzón, M. T., Sandri, L., Rudyy, O., Cheptsov, A., Macedonio, G., Folch, A., Barsotti, S., Selva, J., and Costa, A.: On the feasibility and usefulness of high-performance computing in probabilistic volcanic hazard assessment: An application to tephra hazard from Campi Flegrei, Front. Earth Sci., 10, https://doi.org/10.3389/feart.2022.941789, 2022.

McKay, M. D., Beckman, R. J., and Conover, W. J.: A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis
730 of Output from a Computer Code, Technometrics, 21, 239–245, 1979.

Mingari, L., Folch, A., Dominguez, L., and Bonadonna, C.: Volcanic Ash Resuspension in Patagonia: Numerical Simulations and Observations, Atmosphere, 11, https://doi.org/10.3390/atmos11090977, 2020.

Mingari, L., Folch, A., Prata, A. T., Pardini, F., Macedonio, G., and Costa, A.: Data assimilation of volcanic aerosol observations using FALL3D+PDAF, Atm. Chem. Phys., 22, 1773–1792, https://doi.org/10.5194/acp-22-1773-2022, 2022.

735 O'Neill, N. T., Ignatov, A., Holben, B. N., and Eck, T. F.: The lognormal distribution as a reference for reporting aerosol optical depth statistics; Empirical tests using multi-year, multi-site AERONET Sunphotometer data, Geophys. Res. Lett., 27, 3333–3336, https://doi.org/10.1029/2000GL011581, 2000.

Osores, S., Ruiz, J., Folch, A., and Collini, E.: Volcanic ash forecast using ensemble-based data assimilation: an ensemble transform Kalman filter coupled with the FALL3D-7.2 model (ETKF–FALL3D version 1.0), Geosci. Model Dev., 13, 1–22, https://doi.org/10.5194/gmd-13-
740 1-2020, 2020.

Pardini, F., Corradini, S., Costa, A., Esposti Ongaro, T., Merucci, L., Neri, A., Stelitano, D., et al.: Ensemble-Based Data Assimilation of Volcanic Ash Clouds from Satellite Observations: Application to the 24 December 2018 Mt. Etna Explosive Eruption, Atmosphere, 11, 359, 2020.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V.,
745 Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, J. Mach. Learn. Res., 12, 2825–2830, 2011.

Pelley, R., Cooke, M., Manning, A., Thomson, D., Witham, C., and Hort, M.: Initial implementation of an inversion technique for estimating volcanic ash source parameters in near real time using satellite retrievals, Tech. rep., Tech. rep., Forecasting Research Technical Report, 2015.

750 Pfeiffer, T., Costa, A., and Macedonio, G.: A model for the numerical simulation of tephra fall deposits, J. Volcanol. Geoth. Res., 140, 273 – 294, https://doi.org/10.1016/j.jvolgeores.2004.09.001, 2005.

Posselt, D. J. and Bishop, C. H.: Nonlinear data assimilation for clouds and precipitation using a gamma inverse-gamma ensemble filter, Q. J. Roy. Meteor. Soc., 144, 2331–2349, https://doi.org/10.1002/qj.3374, 2018.

Prata, A. T., Mingari, L., Folch, A., Macedonio, G., and Costa, A.: FALL3D-8.0: a computational model for atmospheric transport and deposition of particles, aerosols and radionuclides – Part 2: Model validation, Geosci. Model Dev., 14, 409–436, https://doi.org/10.5194/gmd-14-409-2021, 2021.

Pyle, D. M.: The thickness, volume and grainsize of tephra fall deposits, Bull. Volcanol., 51, 1–15, 1989.

Romero, J., Morgavi, D., Arzilli, F., Daga, R., Caselli, A., Reckziegel, F., Viramonte, J., Díaz-Alvarado, J., Polacci, M., Burton, M., and Perugini, D.: Eruption dynamics of the 22–23 April 2015 Calbuco Volcano (Southern Chile): Analyses of tephra fall deposits, J. Volcanol. Geoth. Res., 317, 15 – 29, https://doi.org/10.1016/j.jvolgeores.2016.02.027, 2016.

Sandri, L., Costa, A., Selva, J., Tonini, R., Macedonio, G., Folch, A., and Sulpizio, R.: Beyond eruptive scenarios: Assessing tephra fallout hazard from Neapolitan volcanoes, Sci. Rep., 6, 1–13, https://doi.org/10.1038/srep24271, 2016.

Scasso, R. A., Corbella, H., and Tiberi, P.: Sedimentological analysis of the tephra from the 12–15 August 1991 eruption of Hudson volcano, Bull. Volcanol., 56, 121–132, 1994.

Sha, F., Lin, Y., Saul, L. K., and Lee, D. D.: Multiplicative Updates for Nonnegative Quadratic Programming, Neural Comput., 19, 2004–2031, https://doi.org/10.1162/neco.2007.19.8.2004, 2007.

Stefanescu, E. R., Patra, A. K., Bursik, M. I., Madankan, R., Pouget, S., Jones, M., Singla, P., Singh, T., Pitman, E. B., Pavolonis, M., Morton, D., Webley, P., and Dehn, J.: Temporal, probabilistic mapping of ash clouds using wind field stochastic variability and uncertain eruption source parameters: Example of the 14 April 2010 Eyjafjallajökull eruption, J. Adv. Model. Earth Sy., 6, 1173–1184, https://doi.org/10.1002/2014MS000332, 2014.

Suzuki, T.: A theoretical model for dispersion of tephra, in: Arc Volcanism: Physics and Tectonics, edited by Shimozuru, D. and Yokoyama, I., pp. 93–113, Terra Scientific Publishing Company (TERRAPUB), Tokyo, 1983.

Van Eaton, A. R., Amigo, A., Bertin, D., Mastin, L. G., Giacosa, R. E., González, J., Valderrama, O., Fontijn, K., and Behnke, S. A.: Volcanic lightning and plume behavior reveal evolving hazards during the April 2015 eruption of Calbuco volcano, Chile, Geophysical Research Letters, 43, 3563–3571, https://doi.org/10.1002/2016GL068076, 2016.

Vogel, H., Förstner, J., Vogel, B., Hanisch, T., Mühr, B., Schättler, U., and Schad, T.: Time-lagged ensemble simulations of the dispersion of the Eyjafjallajökull plume over Europe with COSMO-ART, Atm. Chem. Phys., 14, 7837–7845, https://doi.org/10.5194/acp-14-7837-2014, 2014.

Wilson, G., Wilson, T., Deligne, N., and Cole, J.: Volcanic hazard impacts to critical infrastructure: A review, J. Volcanol. Geoth. Res., 286, 148–182, https://doi.org/10.1016/j.jvolgeores.2014.08.030, 2014.

Yoshida, T. and Kalnay, E.: Correlation-Cutoff Method for Covariance Localization in Strongly Coupled Data Assimilation, Mon. Weather Rev., 146, 2881 – 2889, https://doi.org/10.1175/MWR-D-17-0365.1, 2018.

Zidikheri, M. J. and Lucas, C.: A Computationally Efficient Ensemble Filtering Scheme for Quantitative Volcanic Ash Forecasts, J. Geophys. Res.-Atmos., 126, e2020JD033 094, https://doi.org/10.1029/2020JD033094, 2021a.

Zidikheri, M. J. and Lucas, C.: Improving Ensemble Volcanic Ash Forecasts by Direct Insertion of Satellite Data and Ensemble Filtering, Atmosphere, 12, https://doi.org/10.3390/atmos12091215, 2021b.

Zidikheri, M. J., Lucas, C., and Potts, R. J.: Estimation of optimal dispersion model source parameters using satellite detections of volcanic ash, J. Geophys. Res.-Atmos., 122, 8207–8232, https://doi.org/10.1002/2017JD026676, 2017.

Zidikheri, M. J., Lucas, C., and Potts, R. J.: Quantitative Verification and Calibration of Volcanic Ash Ensemble Forecasts Using Satellite Data, J. Geophys. Res.-Atmos., 123, 4135–4156, https://doi.org/10.1002/2017JD027740, 2018.

790